进化算法首达时间分析的停时理论模型

张宇山[1),2)] 郝志峰[3)] 黄 翰[4)] 林智勇[5)]

[1)](广东财经大学数学与统计学院 广州 510320)

[2)](南京大学计算机软件新技术国家重点实验室 南京 210093)

[3)](广东工业大学计算机学院 广州 510006)

[4)](华南理工大学软件学院 广州 510006)

[5)](广东技术师范学院计算机科学学院 广州 510665)

摘 要 计算时间分析是进化算法理论基础研究中的重要课题,也是一大难点.该文基于停时理论,结合时齐马氏过程的性质,将进化算法的首达时间视为停时,提出了分析进化算法首达时间的一个新方法.在此框架下,Level-reaching Estimation Technique作为特例得到了严格的证明.为展示如何用该理论方法分析具体问题,以(1+λ)EA求解PEAK函数和(1+λ)ES求解倾斜平面问题为实例,分析了平均首达时间.结果表明,该文所提出的方法不但适用于离散优化问题也适用于连续优化问题,具有通用性.

关键词 进化算法;计算时间;停时;时齐马氏过程;首达时间

中图法分类号 TP18 **DOI号** 10.11897/SP.J.1016.2015.01582

A Stopping Time Theory Model of First Hitting Time Analysis of Evolutionary Algorithms

ZHANG Yu-Shan[1),2)] HAO Zhi-Feng[3)] HUANG Han[4)] LIN Zhi-Yong[5)]

[1)](*School of Mathematics & Statistics, Guangdong University of Finance & Economics, Guangzhou* 510320)

[2)](*State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing* 210093)

[3)](*Faculty of Computer, Guangdong University of Technology, Guangzhou* 510006)

[4)](*School of Software Engineering, South China University of Technology, Guangzhou* 510006)

[5)](*School of Computer Science, Guangdong Polytechnic Normal University, Guangzhou* 510665)

Abstract In contrast to fruitful research results of evolutionary algorithms in practical applications, the theoretical results are still relatively few. Computational time analysis is an important and hard topic in the research of theoretical foundation of evolutionary algorithms. Based on stopping time theory, combining the properties of homogeneous Markov chain, this article regards the first hitting time of evolutionary algorithms as a stopping time and proposes a new general analytic framework. Under this framework, the Level-reaching Estimation Technique is proven rigorously as a special case. To illustrate how the proposed method can be applied to concrete problems in analyzing the expected first hitting time of EAs, we analyze the runtime of (1+λ)EA on PEAK function and (1+λ)ES on inclined plane problem. The results show that the proposed method has generality, it is not only suitable for discrete optimization but also suitable for continuous optimization.

Keywords evolutionary algorithms; computational time; stopping time; homogeneous Markov chain; first hitting time

---

收稿日期:2013-07-18;最终修改稿收到日期:2014-12-12.本课题得到教育部人文社会科学研究青年基金(14YJCZH216)、国家自然科学基金(61370177,61202453)资助.**张宇山**,男,1975年生,博士,副教授,主要研究方向为进化算法的理论基础.E-mail:scuthill@163.com. **郝志峰**,男,1968年生,博士,教授,主要研究领域为算法设计与分析、代数学及组合优化. **黄 翰**,男,1980年生,博士,副教授,主要研究方向为进化计算方法的理论基础、进化计算方法的优化设计及其应用. **林智勇**,男,1977年生,博士,副教授,主要研究方向为算法设计与分析、机器学习和计算智能.

1 引 言

进化算法(Evolutionary Algorithms，EAs)主要包括遗传算法(Genetic Algorithm，GA)、进化规划(Evolutionary Programming，EP)、进化策略(Evolution Strategy，ES)，是受自然进化过程启发而产生的一大类随机启发式算法.遗传算法主要用于求解离散搜索空间的组合优化问题，而进化规划和进化策略则主要针对求解连续优化问题.随着研究的深入，三者有互相融合的趋势.当前，进化算法取得了丰硕的应用成果，主要集中在算法的改进和应用上，大多数是实证研究，相对来说，对于进化算法的运行理论机理研究比较少，这主要归结于进化算法运行的随机性本质.进化算法的理论基础研究有助于回答进化算法“为何有效”及“何时有效”的本质性问题，揭示算法的动态行为，从而对参数的选择和算法的应用具有指导作用.计算时间分析是进化算法理论基础研究中的一大难点和热点[1]，近年来受到国内外相关学者的关注.

进化算法的计算时间，简而言之是指算法找到最优解或满意解所需的迭代次数，亦称运行时间(runtime)，是衡量算法优劣的重要指标.为了便于分析，很多研究都是针对(1+1)EA展开的.尽管(1+1)EA是一种简单的进化算法，但其依然有理论研究的价值，是进化算法计算时间分析的一个基本模型.通过深入研究(1+1)EA，可以揭示进化算法的动态行为，为分析更复杂的算法提供有用的工具.(1+1)EA的个体通常为二进制位串，但也有不采用二进制位串的变体[2-3]，种群规模为1，在每一步迭代中父代个体通过变异产生子代个体，然后选择二者中较优的个体作为下一步迭代的父代个体.Droste等人[4]通过分析(1+1)EA在诸如OneMax函数、LeadingOnes函数、BinVal函数等伪布尔函数(即 $f:\{0,1\}^n\rightarrow R$)上的计算时间，证明了(1+1)EA求解线性伪布尔函数的平均运行时间为 $\Theta(n\ln n)$；他们还构造了一个单峰函数的实例，证明(1+1)EA在其上的计算时间复杂度为指数式，推翻了一些学者之前的猜想.该论文得到的结果以及提出的数学工具和分析方法都极具启发性.Wegener[5]总结了分析(1+1)EA求解伪布尔函数的时间复杂度的常用数学工具和方法.近年来，对(1+1)EA的计算时间分析开始从伪布尔函数推广到具有现实背景的组合优化问题.Oliveto等人[6]针对顶点覆盖问题(Vertex Cover Problems)分析了(1+1)EA在其上的表现，证明对于图中每个节点的度不大于2的实例，(1+1)EA能够有效地找到最小覆盖.计算唯一输入输出序列(Unique Input Output sequences，UIO)是有限状态机一致性测试中的一个基本且困难的问题，Lehre和Yao[2]挑选了此类问题的几个实例，分析了(1+1)EA的时间复杂性.Zhang等人[3]分析了(1+1)EA求解一个TSP实例和指派问题实例的计算时间复杂度.Borisovsky和Eremeev[7]以(1+1)EA作为基准算法，讨论(1+1)EA在何种条件下优于其他的EAs.

针对一般进化算法的首达时间(First Hitting Time，FHT)分析，He和Yao[8-9]提出了一个基于马尔可夫过程的通用框架，这涉及到复杂的矩阵计算.Jansen等人[10]通过分析(1+λ)EA在OneMax和LeadingOnes等函数上的首达时间，初步探讨了子代种群规模大小对进化算法性能的影响.Yu和Zhou[11]结合吸收态马尔可夫链与收敛速率，提出了一个估计平均首达时间的新方法.黄翰、郝志峰等人[12-13]运用吸收态马尔可夫链模型分析了蚁群算法和进化规划的计算时间复杂度.Jägersküpper[14]把马尔可夫链和Drift analysis[15]结合起来，重新分析了(1+1)EA在线性函数上的运行时间上界，得到较精确的估计值.这些研究成果表明，充分利用进化算法的马尔可夫性是分析进化算法计算时间的一条可行之路.除马尔可夫链之外，Drift analysis[15-16]也是一个强大的数学工具，其思想直观，有望成为进化算法计算时间分析的通用工具.但是实际应用中定理条件较难验证.Oliveto和Witt[17]对Drift analysis进行了简化，得到一个更易于验证的Drift定理，计算了(1+1)EA在几个伪布尔函数和最大匹配问题实例上的运行时间下界.Chen等人[18]对基于种群的EAs在单峰函数上的时间复杂度展开了分析.他们通过将takeover time概念和Drift analysis方法相结合，提出了一个新的分析方法.他们选择两个单峰函数(OneMax和LeadingOnes)为实例，分析了不带交叉算子的($N+N$)EA的平均首达时间上界，并讨论了将此方法推广到处理更复杂问题的可能性.

当前的进化算法计算时间分析主要集中在离散搜索空间的情形，连续型进化算法如进化规划和进化策略的计算时间分析成果目前相对较少.

Jägerskü pper[19-20]对基本的进化策略——(1+1)ES在连续优化中的运行时间进行了研究,其分析过程比较冗长. 黄翰等人[13]运用吸收态马尔可夫链模型,建立了进化规划算法时间复杂度分析的一般框架,并以Gauss变异进化规划算法为例,分析了期望收敛时间. Agapie等人[21]引入更新过程理论,分析了带均匀变异算子的(1+1)ES在倾斜平面(inclined plane)问题上的计算时间,取得了比较精确的结果. 需要指出的是,按更新过程$\{N(t),t\geq 0\}$的定义,$N(t)=\sup\{n:S_n\leq t\}$,$S_0=0$,$S_n=\sum_{k=1}^{n}X_k$,$\{X_k,k\geq 1\}$是独立同分布的非负随机变量. 为了能引用更新过程的有关结论,需假设(1+1)ES在倾斜平面上的每一步位移独立同分布. 这是很强的假设,甚至比马尔可夫性要求还高,现实中只有少数情况才可能满足,所以该分析方法的适用范围受到很大的局限.

受已有的理论分析方法如Drift analysis、马尔可夫过程模型、更新过程模型等的思想的启发,本文引入停时(stopping time)理论作为工具,把进化算法的FHT视作停时,利用时齐马氏过程的性质,提出了分析进化算法平均首达时间(Expected First Hitting Time,EFHT)的通用方法. 与文献[21]所用的更新过程模型的不同之处在于,本文所提方法无需对进化算法的一步位移作独立同分布的假设,即可得到更具一般性的结果. Level-reaching Estimation Technique是离散型进化算法时间复杂度分析中常用的工具[1,3-5,22],本文的分析表明它是我们所提出的定理的一个特例. 最后,用所提出的方法对两个具体实例作了分析:(1)(1+λ)EA优化PEAK函数;(2)(1+λ)ES求解倾斜平面问题. 前者是离散型优化,后者是连续型优化,算法的子代种群规模都大于1,更贴近实际.

2 进化算法与停时

2.1 进化算法的随机过程模型

不失一般性,本文讨论的最优化问题可描述为:设可行解空间(搜索空间)为非空集合S,R表示实数集,$f:S\to R$为目标函数,欲搜索到至少1个全局极大值点$x^*\in S$,使得$f_{\max}=f(x^*)=\max_{x\in S}f(x)$.

进化算法的求解流程大致可用以下的框架统一描述[23].

算法1. 进化算法.

输入:初始种群$p(0)$

输出:包含全局最优解的种群$p(t)$

1. 初始化,生成初始种群$p(0)=\{x_1^0,x_2^0,\cdots,x_N^0\}$,令$t=0$;

2. 计算$p(t)$中每一个个体的适应值以评价当前种群的优劣;

3. 用进化算子生成新一代种群$p(t+1)$,令$t=t+1$;

4. 若终止条件满足,则停止;否则回到步2.

常用的进化算子主要有3种:交叉、变异、选择. 在求解实际问题时,遗传算法一般采用全部3种进化算子,而进化规划和进化策略只采用后面两种. 本文考察的进化算法采用$(\mu+\lambda)$选择机制,即每次迭代时,由μ个父代个体生成λ个子代个体,然后将这$\mu+\lambda$个个体按适应值从大到小排列,选择排在前面的μ个个体作为下一步迭代的父代个体.

进化算法中,搜索空间S的元素x称为一个个体或解,可直接选取目标函数$f:S\to R$作为个体的适应值函数. 若干个个体组成一个种群,一个规模为N的种群可以表示为$X=\{x_1,x_2,\cdots,x_N\}$,$x_i\in S$,$i=1,2,\cdots,N$. 种群空间由全体可能的种群构成,记为$E=\{X=\{x_1,x_2,\cdots,x_N\}\mid x_i\in S,i=1,2,\cdots,N\}$. 针对本文所考察的最大化优化问题,定义$S_{\rm opt}=\{x^*\in S\mid f(x^*)=\max_{x\in S}f(x)\}$为进化算法的最优个体空间,最优种群空间则定义为$E_{\rm opt}=\{X=\{x_1,x_2,\cdots,x_N\}\mid \exists x_i\in X,i=1,2,\cdots,N$,使得$f(x_i)=f_{\max}\}$.

设$X_k=\{x_1^k,x_2^k,\cdots,x_N^k\}$,$k=0,1,\cdots$为进化算法的第$k$代种群,则$\{X_k,k=0,1,\cdots\}$是一个在状态空间$E$上取值的随机过程. 进一步的,由于第$k+1$代种群$X_{k+1}$的状态仅依赖于第$k$代种群$X_k$而不受之前种群的影响,且本文考察的进化算法不采用自适应变异机制,因此随机过程$\{X_k,k=0,1,\cdots\}$可视为时齐马尔可夫过程[9,15,24]. 为衡量某个种群与最优种群空间的距离,我们引入下列的定义.

定义1. 最优个体序列. 令ξ_k为第k代种群X_k的一个个体,即$\xi_k\in X_k$,$k=0,1,\cdots$. 若$f(\xi_k)=\max\{f(x_i^k),i=1,2,\cdots,N\}$,则把随机过程$\{\xi_k,k=0,1,\cdots\}$称为进化算法的最优个体序列.

设$S_{\rm obj}$表示搜索空间S中的某一特定区域,不妨称之为目标区域,距离函数的定义如下.

定义2. 个体距离函数. 令$S_{\rm obj}\subset S$为目标区域,对于定义在搜索空间S的函数$d:S\to R^+$,若其满足:

(1) $d(x)=0\Leftrightarrow x\in S_{\mathrm{obj}}$；

(2) 对于 $\forall x,y\in S\setminus S_{\mathrm{obj}}$，有 $d(x)=d(y)\Leftrightarrow f(x)=f(y)$；

(3) 对于 $\forall x,y\in S\setminus S_{\mathrm{obj}}$，有 $d(x)<d(y)\Leftrightarrow f(x)>f(y)$.

则将 $d:S\to R^+$ 称为个体到目标区域的距离函数. 特别的，若 $S_{\mathrm{obj}}=S_{\mathrm{opt}}$，则 d 简称为距离函数.

定义 2 所提的距离函数是存在的，如可令$d(x)=f_{\max}-f(x)$，$x\in S$，即为满足定义 2 的一种距离函数.

对于种群 $X=\{x_1,x_2,\cdots,x_N\}$，$x_i\in S$，$i=1,2,\cdots,N$，把 $V(X)=d(\tilde{x})$称为种群 X 的距离函数，其中 $\tilde{x}\in X$ 满足 $f(\tilde{x})=\max\limits_{x\in X}f(x)$，即种群的距离用其中最优个体的距离来定义.

当进化算法使用$(\mu+\lambda)$选择机制的精英选择策略时，显然 $V(X_k)=d(\xi_k)$，$k=0,1,\cdots$是一个单调不增的序列. $d(\xi_k)$越小，表明种群与最优种群空间的距离越小，即越趋近于最优解，当 $d(\xi_k)=0$ 时，则说明算法已经求得最优解. 在连续优化中要求 $d(\xi_k)=0$ 并不现实，有鉴于此，我们引入 ε-FHT 的定义.

定义 3.　ε-FHT. 假定随机过程$\{\xi_k,k=0,1,\cdots\}$为进化算法的某个最优个体序列，实数 $\varepsilon>0$，定义$\tau_\varepsilon=\min\{k\geq 0;d(\xi_k)<\varepsilon\}$为进化算法的 ε-FHT.

ε-FHT 刻画了在进化过程中种群与最优种群空间的距离首次小于 ε 时，算法所需要迭代的次数.

接下来，本文将采用停时理论为工具，研究进化算法的 ε-FHT.

2.2　停时理论简介

停时作为随机过程理论中的一个重要概念，从直观上来讲是一个随机时间. 其严格定义如下[25].

定义 4.　停时. 设 T 为非负整值随机变量，$\{Y_n,n\geq 0\}$为一随机序列，令$\mathcal{F}_n=\sigma(Y_k;0\leq k\leq n)$.

若对 $\forall n\geq 0$，有$\{T=n\}\in\mathcal{F}_n$，则称 T 是$\{Y_n,n\geq 0\}$的停时.

在上述定义中，称$\mathcal{F}_n=\sigma(Y_k;0\leq k\leq n)=\sigma(Y_0,Y_1,\cdots,Y_n)$为由随机变量序列 $Y_0,Y_1,\cdots,Y_n$生成的事件σ 域，它代表由 $Y_0,Y_1,\cdots,Y_n$生成的全部事件信息. 显然，$\mathcal{F}_0\subset\mathcal{F}_1\subset\cdots\subset\mathcal{F}_n\subset\cdots$. 直接使用定义 4 有时不太方便，为此，我们推出如下的等价定义.

推论 1.　随机变量 T 是$\{Y_n,n\geq 0\}$的停时，当且仅当对 $\forall n\geq 0$，有$\{T\leq n\}\in\mathcal{F}_n$.

证明.　充分性.

若对 $\forall n\geq 0$，有$\{T\leq n\}\in\mathcal{F}_n$，因为$\{T=n\}=\{T\leq n\}\setminus\{T\leq n-1\}$，已知$\{T\leq n\}\in\mathcal{F}_n$，$\{T\leq n-1\}\in\mathcal{F}_{n-1}\subset\mathcal{F}_n$，由$\mathcal{F}_n$是一个 σ 域，可得$\{T=n\}=\{T\leq n\}\setminus\{T\leq n-1\}\in\mathcal{F}_n$，即 T 是$\{Y_n,n\geq 0\}$的停时.

必要性.

若T是$\{Y_n,n\geq 0\}$的停时，即 $\forall n\geq 0$，$\{T=n\}\in\mathcal{F}_n$，此时$\{T\leq n\}=\bigcup\limits_{k=0}^{n}\{T=k\}$，而当 $0\leq k\leq n$ 时，$\{T=k\}\in\mathcal{F}_k\subset\mathcal{F}_n$，由于$\mathcal{F}_n$是一个 σ 域，故$\{T\leq n\}\in\mathcal{F}_n$.

证毕.

进化算法首次找到满意解的时间仅仅与当前以及过去的搜索结果有关，故有以下引理.

引理 1.　进化算法的 ε-FHT 是种群序列$\{X_k,k=0,1,\cdots\}$的停时.

证明.　由定义 3，对 $\forall n\geq 0$，$\{\tau_\varepsilon=n\}=\{d(\xi_k)\geq\varepsilon,k=0,\cdots,n-1,d(\xi_n)<\varepsilon\}=\bigcap\limits_{k=0}^{n-1}\{d(\xi_k)\geq\varepsilon\}\cap\{d(\xi_n)<\varepsilon\}\in\sigma(X_k;0\leq k\leq n)$.

根据定义 4，结论得证.

证毕.

3　进化算法的平均首达时间分析

对$\mathbb{E}(\tau_\varepsilon)$的上下界进行分析之前，需将 τ_ε的原始表达式作一番转换.

令 $\eta_t=d(\xi_{t-1})-d(\xi_t)$，$t=1,2,\cdots$代表迭代过程中进化算法的种群从第 $t-1$ 步到第 t 步的一步位移，此时

$$\begin{aligned}d(\xi_k)<\varepsilon&\Leftrightarrow d(\xi_0)-d(\xi_k)>d(\xi_0)-\varepsilon\\&\Leftrightarrow d(\xi_0)-d(\xi_1)+d(\xi_1)-d(\xi_2)+\cdots+\\&\quad\ d(\xi_{k-1})-d(\xi_k)>d(\xi_0)-\varepsilon\\&\Leftrightarrow\sum_{t=1}^{k}\eta_t>d(\xi_0)-\varepsilon.\end{aligned}$$

即 $\tau_\varepsilon=\min\{k\geq 0;d(\xi_k)<\varepsilon\}=\min\left\{k\geq 0,\sum\limits_{t=1}^{k}\eta_t>d(\xi_0)-\varepsilon\right\}$，此处约定$\sum\limits_{t=1}^{0}\eta_t=0$.

下文中的 I_A代表事件 A 的示性函数，$I_A=1$ 当且仅当 A 发生，否则 $I_A=0$. 现在我们给出本文的主要定理，考虑到进化算法的首达时间与初始种群的状态有关，为了更精细地刻画首达时间与初始种群的关系，定理 1 给出了$\mathbb{E}[\tau_\varepsilon|X_0]$的估计式.

定理 1.　设$\{X_k,k=0,1,\cdots\}$为进化算法的种群序列，$\{\xi_k,k=0,1,\cdots\}$为对应的最优个体序列. 当 $d(\xi_0)>0$ 时，若存在两个关于 X_0 的函数 $\alpha(X_0)$，$\beta(X_0)$，使 $0<\alpha(X_0)\leq\mathbb{E}(\eta_1|X_0)\leq\beta(X_0)$，则对任意的 $\varepsilon>0$ 有

$$\frac{d(\xi_0)-\varepsilon}{\beta(X_0)} \le \mathbb{E}[\tau_\varepsilon \mid X_0] \le \frac{d(\xi_0)}{\alpha(X_0)} \tag{1}$$

证明. 因为$\sum_{k=1}^{\tau_\varepsilon}\eta_k=\sum_{k=1}^{\infty}\eta_k I_{\{\tau_\varepsilon\ge k\}}$是非负项级数，根据 Levi 单调收敛定理得

$$\mathbb{E}\left(\sum_{k=1}^{\tau_\varepsilon}\eta_k \mid X_0\right)=\mathbb{E}\left(\sum_{k=1}^{\infty}\eta_k I_{\{\tau_\varepsilon\ge k\}} \mid X_0\right)$$

$$=\sum_{k=1}^{\infty}\mathbb{E}[\eta_k I_{\{\tau_\varepsilon\ge k\}} \mid X_0],$$

记$\mathcal{F}_n=\sigma(X_i;0\le i\le n)$，由条件期望的性质可得

$$\text{上式}=\sum_{k=1}^{\infty}\mathbb{E}[\mathbb{E}(\eta_k I_{\{\tau_\varepsilon\ge k\}} \mid \mathcal{F}_{k-1}) \mid X_0].$$

根据引理 1，τ_ε是$\{X_k,k=0,1,\cdots\}$的停时，再由推论 1 可知$\{\tau_\varepsilon\ge k\}=\{\tau_\varepsilon\le k-1\}^c\in\mathcal{F}_{k-1}$，即$I_{\{\tau_\varepsilon\ge k\}}$关于$\mathcal{F}_{k-1}$可测，故

$$\text{上式}=\sum_{k=1}^{\infty}\mathbb{E}[I_{\{\tau_\varepsilon\ge k\}}\ \mathbb{E}(\eta_k \mid \mathcal{F}_{k-1}) \mid X_0].$$

因为$\{X_k,k=0,1,\cdots\}$是一个时齐马氏过程，故当$\tau_\varepsilon\ge k$时，$\mathbb{E}(\eta_k \mid \mathcal{F}_{k-1})=\mathbb{E}(\eta_k \mid X_{k-1})=\mathbb{E}(\eta_1 \mid X_0)$，从而$\alpha(X_0)\le\mathbb{E}(\eta_k \mid \mathcal{F}_{k-1})\le\beta(X_0)$. 于是

$$\sum_{k=1}^{\infty}\mathbb{E}[\alpha(X_0)I_{\{\tau_\varepsilon\ge k\}}|X_0]\le\sum_{k=1}^{\infty}\mathbb{E}[I_{\{\tau_\varepsilon\ge k\}}\mathbb{E}(\eta_k|\mathcal{F}_{k-1})|X_0]$$

$$\le\sum_{k=1}^{\infty}\mathbb{E}[\beta(X_0)I_{\{\tau_\varepsilon\ge k\}} \mid X_0].$$

注意到$\alpha(X_0)$，$\beta(X_0)$是关于X_0的函数，马上得到

$$\alpha(X_0)\sum_{k=1}^{\infty}\mathbb{E}(I_{\{\tau_\varepsilon\ge k\}} \mid X_0)\le\mathbb{E}\left(\sum_{k=1}^{\tau_\varepsilon}\eta_k \mid X_0\right)$$

$$\le\beta(X_0)\sum_{k=1}^{\infty}\mathbb{E}(I_{\{\tau_\varepsilon\ge k\}} \mid X_0),$$

而$\sum_{k=1}^{\infty}\mathbb{E}(I_{\{\tau_\varepsilon\ge k\}}|X_0)=\sum_{k=1}^{\infty}\mathbb{P}(\tau_\varepsilon\ge k|X_0)=\mathbb{E}(\tau_\varepsilon|X_0)$[25].

显然$d(\xi_0)-\varepsilon<\sum_{k=1}^{\tau_\varepsilon}\eta_k\le d(\xi_0)$，$d(\xi_0)$也是关于$X_0$的函数，故

$$d(\xi_0)-\varepsilon\le\mathbb{E}\left(\sum_{k=1}^{\tau_\varepsilon}\eta_k \mid X_0\right)\le d(\xi_0).$$

最后得到结论

$$\frac{d(\xi_0)-\varepsilon}{\beta(X_0)} \le \mathbb{E}[\tau_\varepsilon \mid X_0] \le \frac{d(\xi_0)}{\alpha(X_0)}.$$

证毕.

估算种群的一步平均位移$\mathbb{E}(\eta_1 \mid X_0)$是定理 1 应用的关键，作为条件数学期望，它是关于X_0的一个函数. 为使计算更为简便，可选择常数作为$\mathbb{E}(\eta_1 \mid X_0)$的上下界，但这有可能使得下界过小、上界过大，从而导致对$\mathbb{E}[\tau_\varepsilon \mid X_0]$的上下界估计过于宽松. 为此定理 1 采用两个关于X_0的函数$\alpha(X_0)$，$\beta(X_0)$作为$\mathbb{E}(\eta_1 \mid X_0)$的下上界，有助于更精细地估计$\mathbb{E}[\tau_\varepsilon \mid X_0]$的上下界.

推论 2 综合考虑初始种群的各种状态，可以得到无条件数学期望$\mathbb{E}[\tau_\varepsilon]$的上下界.

推论 2. 若定理 1 的条件满足，则有

$$\mathbb{E}\left[\frac{d(\xi_0)-\varepsilon}{\beta(X_0)}\right]\le\mathbb{E}[\tau_\varepsilon]\le\mathbb{E}\left[\frac{d(\xi_0)}{\alpha(X_0)}\right].$$

特别地，当$\alpha(X_0)$，$\beta(X_0)$分别为两个常数C_{low}，C_{up}时，有$\frac{\mathbb{E}[d(\xi_0)]-\varepsilon}{C_{\text{up}}}\le\mathbb{E}[\tau_\varepsilon]\le\frac{\mathbb{E}[d(\xi_0)]}{C_{\text{low}}}$.

证明. 由$\mathbb{E}[\tau_\varepsilon]=\mathbb{E}[\mathbb{E}[\tau_\varepsilon \mid X_0]]$立即得证.

证毕.

Level-reaching Estimation Technique 亦称 Method of Fitness-Based Partitions，该方法主要用于分析采用精英选择策略的种群规模为 1 的进化算法，是一个简单而实用的数学工具. 该方法首先按照适应值函数的取值对搜索空间作适当的分割，然后估计个体从当前区域跳到更优区域的概率下界，从而估算出算法最终找到最优解的计算时间. 该方法原有的理论证明[5,26]参考了几何分布的性质，直观但不够严格. 通过构造适当的距离函数，在本文所建立的理论框架下，该方法实际是定理 1 的直接推论，可得到严格的证明.

定义 5. 基于f的分割. 设$f:S\to R$是待极大化的目标函数，S是非空有限集. 把f全部可能取到的值按升序排列$f_0<f_1<\cdots<f_M=f_{\max}$. 令$A_i=\{x\in S|f(x)=f_i\}$，$i=0,1,\cdots,M$，则称$\{A_0,A_1,\cdots,A_M\}$是基于$f$的分割.

对于采用精英选择策略的种群规模为 1 的进化算法，最优个体序列$\{\xi_k,k=0,1,\cdots\}$即为种群序列. Level-reaching Estimation Technique 的严格证明如下.

推论 3. 设$\{A_0,A_1,\cdots,A_M\}$是一个基于f的分割，记$s(a)=\mathbb{P}\{\xi_{t+1}=a'\in A_{i+1}\cup\cdots\cup A_M \mid \xi_t=a\in A_i\}$，$t=0,1,\cdots$；$i=0,1,\cdots,M-1$. 令$s_i=\min\{s(a) \mid a\in A_i\}$，$\tau=\min\{k\ge 0;\xi_k\in S_{\text{opt}}\}$，则可得

$$\mathbb{E}[\tau]\le\sum_{i=0}^{M-1}s_i^{-1} \tag{2}$$

证明. 考察算法的搜索过程，不妨记$\tau_i=\min\{k\ge 1;\xi_k\in A_{i+1}\cup\cdots\cup A_M,\xi_0\in A_i\}$为个体从进入$A_i$，$i=0,1,\cdots,M-1$直到离开为止的逗留时间. 显然$\tau\le\sum_{i=0}^{M-1}\tau_i$. 为方便表述，下文中的下标$i$约定总

取 $i=0,1,\cdots,M-1$.

在这里 $S_{\mathrm{obj}}=A_{i+1}\cup\cdots\cup A_M$，令 $d(x)=\begin{cases}1,\ x\in A_i\\0,\ x\in A_{i+1}\cup\cdots\cup A_M\end{cases}$，当 $d(\xi_0)>0$ 时，根据已知条件可得

$$\begin{aligned}\mathbb{E}(\eta_1|X_0)&=\mathbb{E}(\eta_1|\xi_0)\\&=0\cdot\mathbb{P}(\eta_1=0|\xi_0)+1\cdot\mathbb{P}(\eta_1=1|\xi_0)\geq s_i.\end{aligned}$$

根据推论 2，得 $\mathbb{E}[\tau_i]\leq\dfrac{\mathbb{E}[d(\xi_0)]}{s_i}=s_i^{-1}$，从而 $\mathbb{E}[\tau]\leq\mathbb{E}\left[\sum_{i=0}^{M-1}\tau_i\right]\leq\sum_{i=0}^{M-1}s_i^{-1}$.

证毕.

精确地估计一步转移概率的下界 s_i，$i=0,1,\cdots,M-1$ 是应用推论 3 的关键.

4　实例分析

4.1　(1+λ)EA 求解 PEAK 函数

PEAK 函数只在一处取得最大值，其他的位串则具有相同的适应值，从而形成一个适应值平台，因此也被称为"大海捞针"(Needle-in-the-haystack)函数. 由于进化算法的搜索过程受适应值引导，因此从直观上看，算法很难逃离该平台. 针对只接受具有更高适应值的个体的(1+1)EA，Droste 等人[4]分析了其在 PEAK 函数上的 EFHT，得到了 $\Omega\left(\left(\dfrac{n}{2}\right)^n\right)$ 的下界. 当采用规模大于 1 的子代种群时，算法的性能有否提升是一个值得探讨的问题. 本小节将运用定理 1 分析(1+λ)EA 求解 PEAK 函数的平均首达时间下界. 待研究的目标函数是

$$\mathrm{PEAK}(\boldsymbol{x})=\prod_{i=1}^{n}x_i,\ \boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in\{0,1\}^n,$$

显然，该函数在(1,1,…,1)处取到最大值 1，而在其他地方的取值都为 0.

为方便作比较，本小节所讨论的(1+λ)EA 仅仅接受适应值较大的个体作为下一次迭代的父代个体，算法流程如下所示.

算法 2.　(1+λ)EA.

输入：初始个体

输出：最优个体

1. 初始化，随机生成 1 个个体 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in\{0,1\}^n$；
2. 通过变异产生 λ 个中间个体 $\boldsymbol{y}^1,\boldsymbol{y}^2,\cdots,\boldsymbol{y}^\lambda$；
3. 若 $m=\max\{f(\boldsymbol{y}^1),f(\boldsymbol{y}^2),\cdots,f(\boldsymbol{y}^\lambda)\}>f(\boldsymbol{x})$，则从 $\{\boldsymbol{y}^i|1\leq i\leq\lambda$ 且 $f(\boldsymbol{y}^i)=m\}$ 中任选一个取代 $\boldsymbol{x}$，否则，保留 $\boldsymbol{x}$；
4. 若满足终止条件，则停止，否则回到步 2.

步 2 采用逐位翻转(bit-wise flip)变异：对于父代个体 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 中的每一位 x_i，$i=1,2,\cdots,n$，以 $1/n$ 的概率独立翻转，产生一个中间个体 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$，此时 $\mathbb{P}\{y_i=1-x_i\}=1/n$. 对父代个体 $\boldsymbol{x}$ 作 λ 次重复变异操作，共产生 λ 个独立的中间个体.

记 $\{\boldsymbol{\xi}_k,k=0,1,\cdots\}$ 为(1+λ)EA 的个体序列，$\tau=\min\{k\geq0;\boldsymbol{\xi}_k=(1,1,\cdots,1)\}$ 为算法在这个问题上的首达时间.

定理 2.　(1+λ)EA 在 PEAK 函数上的平均首达时间是 $\mathbb{E}[\tau]=\Omega\left(\dfrac{1}{\lambda}\left(\dfrac{n}{2}\right)^n\right)$.

证明.　首先给出距离函数的定义，令 $\boldsymbol{\xi}$ 为算法产生的任一个个体，定义

$$d(\boldsymbol{\xi})=\begin{cases}0,&\boldsymbol{\xi}=(1,1,\cdots,1);\\1,&\boldsymbol{\xi}\in\{0,1\}^n\backslash\{(1,1,\cdots,1)\}.\end{cases}$$

令 $\eta_t=d(\boldsymbol{\xi}_{t-1})-d(\boldsymbol{\xi}_t)$，$t=1,2,\cdots$，下面估计 $\mathbb{E}(\eta_1|\boldsymbol{\xi}_0)$. 假设 $\boldsymbol{\xi}_0$ 中包含 $i\geq1$ 个 0，由于 1 个父代个体通过 λ 次独立变异，总共产生 λ 个中间个体，于是

$$\begin{aligned}\mathbb{E}(\eta_1|\boldsymbol{\xi}_0)&=\mathbb{P}\{\boldsymbol{\xi}_1=(1,1,\cdots,1)|\boldsymbol{\xi}_0\}\\&=1-\left(1-\left(\frac{1}{n}\right)^i\left(1-\frac{1}{n}\right)^{n-i}\right)^\lambda.\end{aligned}$$

根据伯努利不等式：对任意整数 $n\geq0$，和任意实数 $x>-1$，有 $(1+x)^n\geq1+nx$ 成立，于是

$$\begin{aligned}\mathbb{E}(\eta_1|\boldsymbol{\xi}_0)&=1-\left(1-\left(\frac{1}{n}\right)^i\left(1-\frac{1}{n}\right)^{n-i}\right)^\lambda\\&\leq\lambda\left(\frac{1}{n}\right)^i\left(1-\frac{1}{n}\right)^{n-i}.\end{aligned}$$

记 $\beta(\boldsymbol{\xi}_0)=\lambda\left(\dfrac{1}{n}\right)^i\left(1-\dfrac{1}{n}\right)^{n-i}$，由定理 1 可知

$$\mathbb{E}[\tau|\boldsymbol{\xi}_0]\geq\frac{d(\boldsymbol{\xi}_0)}{\beta(\boldsymbol{\xi}_0)}=\frac{1}{\lambda}n^i\left(\frac{n}{n-1}\right)^{n-i}.$$

再根据推论 2 得

$$\begin{aligned}\mathbb{E}[\tau]&=\mathbb{E}[\mathbb{E}[\tau|\boldsymbol{\xi}_0]]\\&=\sum_{\boldsymbol{x}\neq(1,1,\cdots,1)}\mathbb{E}[\tau|\boldsymbol{\xi}_0]\,\mathbb{P}\{\boldsymbol{\xi}_0=\boldsymbol{x}\}+\\&\quad\ \mathbb{E}[\tau|\boldsymbol{\xi}_0]\,\mathbb{P}\{\boldsymbol{\xi}_0=(1,1,\cdots,1)\}\\&\geq\sum_{i=1}^{n}\frac{1}{\lambda}n^i\left(\frac{n}{n-1}\right)^{n-i}\binom{n}{i}\left(\frac{1}{2}\right)^n\\&=\frac{1}{\lambda}\left(\frac{1}{2}\right)^n\sum_{i=1}^{n}\binom{n}{i}n^i\left(\frac{n}{n-1}\right)^{n-i}\\&=\frac{1}{\lambda}\left(\frac{1}{2}\right)^n\left(\sum_{i=0}^{n}\binom{n}{i}n^i\left(\frac{n}{n-1}\right)^{n-i}-\left(\frac{n}{n-1}\right)^n\right)\end{aligned}$$

$$=\frac{1}{\lambda}\left(\frac{1}{2}\right)^{n}\left(\left(n+\frac{n}{n-1}\right)^{n}-\left(\frac{n}{n-1}\right)^{n}\right)$$

$$\geq\frac{1}{\lambda}\left(\frac{1}{2}\right)^{n}n^{n}=\frac{1}{\lambda}\left(\frac{n}{2}\right)^{n}.$$

即
$$\mathbb{E}[\tau]=\Omega\left(\frac{1}{\lambda}\left(\frac{n}{2}\right)^{n}\right).$$

证毕.

因为(1+λ)EA 在每次迭代时需要对 λ 个中间个体的适应值进行评估,所以在首次找到最优解前算法需要的适应值评估次数平均为 $\lambda\mathbb{E}[\tau]=\Omega\left(\left(\frac{n}{2}\right)^{n}\right)$. 由此可见,增大子代种群的规模并没有降低算法求解 PEAK 函数的计算复杂度.

4.2 (1+λ)ES 求解倾斜平面问题

倾斜平面(inclined plane)问题是一个基本的连续优化问题,Agapie 等人[21]引入更新过程理论,分析了带均匀分布变异算子的(1+1)ES 在该问题上的平均运行时间. 本小节应用定理 1 将他们的分析推广到带正态分布变异算子的(1+λ)ES. 本节讨论的目标函数是二维倾斜平面(inclined plane)[20]:

$f(x_1,x_2)=x_1,(x_1,x_2)\in S=[0,a]\times[-a,a],a>0.$

显然,$\max\limits_{(x_1,x_2)\in S}f(x_1,x_2)=a$,当 $x_1=a$ 时取到最大值.

(1+λ)ES 的算法流程如算法 3 所示.

算法 3. (1+λ)ES.

输入:初始个体

输出:最优个体

1. 初始化,随机生成 1 个个体 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in R^n$;
2. 通过变异产生 λ 个中间个体;
3. 从这 $1+\lambda$ 个个体中选择 1 个最佳个体;
4. 若终止条件满足,则停止;否则回到步 2.

步 2 的变异算子采用标准正态分布:对于父代个体 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 中的每个分量 $x_i,i=1,2,\cdots,n$,令 $y_i=x_i+z_i,z_i\sim N(0,1),z_1,z_2,\cdots,z_n$ 相互独立,生成一个中间个体 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$. 对父代个体 $\boldsymbol{x}$ 重复 λ 次变异操作,共独立产生 λ 个中间个体.

由于本小节讨论的是二维连续优化问题,故可设 $\boldsymbol{\xi}_t=(x_1^t,x_2^t)\in S\subset R^2,t=0,1,\cdots$ 为算法的第 t 代个体. 距离函数定义为 $d(\boldsymbol{\xi}_t)=a-f(\boldsymbol{\xi}_t)=a-x_1^t$,则 $\eta_t=d(\boldsymbol{\xi}_{t-1})-d(\boldsymbol{\xi}_t)=x_1^t-x_1^{t-1},t=1,2,\cdots$ 为算法的个体在迭代过程中从第 $t-1$ 步到第 t 步的位移. 令 $\tau_\varepsilon=\min\{k\geq0;d(\boldsymbol{\xi}_k)<\varepsilon\}=\min\{k\geq0;x_1^k>a-\varepsilon\}$,则 τ_ε 为(1+λ)ES 在倾斜平面问题上的 ε-FHT.

下面的引理描述了 $\eta_t,t=1,2,\cdots$ 的分布:

引理 2. $\eta_t,t=1,2,\cdots$ 独立同分布,分布函数 $F(x)=\mathbb{P}\{\eta_t\leq x\}$ 为

$$F(x)=\begin{cases}0, & x<0\\ \left(\frac{1}{2}\right)^{\lambda}, & x=0\\ \left(\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v\right)^{\lambda}, & x>0\end{cases}\tag{3}$$

证明. 记 $x_1^t,t=0,1,\cdots$ 表示第 t 代个体的第 1 个分量,$\tilde{x}_{1,1}^t,\tilde{x}_{1,2}^t,\cdots,\tilde{x}_{1,\lambda}^t$ 分别为第 t 代个体经变异后产生的 λ 个中间个体的第 1 个分量,则 $\tilde{x}_{1,k}^t=x_1^t+z_k,z_k\sim N(0,1),k=1,2,\cdots,\lambda$ 且相互独立. 于是第 $t+1$ 代个体的第一个分量为

$$x_1^{t+1}=\begin{cases}x_1^t, & \max\limits_{k=1,2,\cdots,\lambda}\tilde{x}_{1,k}^t\leq x_1^t\\ \max\limits_{k=1,2,\cdots,\lambda}\tilde{x}_{1,k}^t, & \max\limits_{k=1,2,\cdots,\lambda}\tilde{x}_{1,k}^t>x_1^t\end{cases}.$$

根据 $\eta_t,t=1,2,\cdots$ 的定义,可得

$$\eta_t=\begin{cases}0, & \max\limits_{k=1,2,\cdots,\lambda}z_k\leq0,\\ \max\limits_{k=1,2,\cdots,\lambda}z_k, & 0<\max\limits_{k=1,2,\cdots,\lambda}z_k.\end{cases}\tag{4}$$

由算法的搜索过程可知,每一次迭代时正态随机变量 $z_k,k=1,2,\cdots,\lambda$ 都是独立产生的,故 $\eta_t,t=1,2,\cdots$ 独立同分布,其分布函数 $F(x)=\mathbb{P}\{\eta_t\leq x\}$ 为

(1) 当 $x<0$ 时:$F(x)=0$;

(2) 当 $x=0$ 时:

$$F(x)=\mathbb{P}\{\eta_t\leq x\}=\mathbb{P}\{\eta_t=0\}=\mathbb{P}\{\max_{k=1,\cdots,\lambda}z_k\leq0\}$$

$$=\prod_{k=1}^{\lambda}\mathbb{P}\{z_k\leq0\}=\left(\frac{1}{2}\right)^{\lambda};$$

(3) 当 $x>0$ 时:

$$F(x)=\mathbb{P}\{\eta_t\leq x\}=\mathbb{P}\{\max_{k=1,\cdots,\lambda}z_k\leq x\}$$

$$=\prod_{k=1}^{\lambda}\mathbb{P}\{z_k\leq x\}=\left(\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v\right)^{\lambda}.$$

即
$$F(x)=\begin{cases}0, & x<0,\\ \left(\frac{1}{2}\right)^{\lambda}, & x=0,\\ \left(\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v\right)^{\lambda}, & x>0.\end{cases}$$

证毕.

不妨假定算法在初始化时从原点出发,即 $\boldsymbol{\xi}_0=(x_1^0,x_2^0)=(0,0)$,此时有如下结论.

定理 3. 带标准正态变异算子的(1+λ)ES 求解倾斜平面问题的平均 ε-FHT 满足

$$\frac{\sqrt{2\pi}(a-\varepsilon)}{\lambda}\leq\mathbb{E}[\tau_\varepsilon|\boldsymbol{\xi}_0=(0,0)]\leq\frac{\sqrt{2\pi}a}{\lambda\left(\frac{1}{2}\right)^{\lambda-1}}\tag{5}$$

证明.　由式(4)可知，η_1只与标准正态随机变量z_k，$k=1,2,\cdots,\lambda$有关，故η_1与$\boldsymbol{\xi}_0$相互独立. 于是根据式(3)可得

$$\mathbb{E}(\eta_1|\boldsymbol{\xi}_0)=\mathbb{E}(\eta_1)=\int_{-\infty}^{+\infty}x\mathrm{d}F(x)=$$
$$\int_0^{+\infty}x\cdot\lambda\left(\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v\right)^{\lambda-1}\cdot\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x.$$

而当$x>0$时，$\frac{1}{2}\le\left(\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v\right)\le 1$，故

$$\lambda\left(\frac{1}{2}\right)^{\lambda-1}\int_0^{+\infty}x\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x\le\mathbb{E}(\eta_1|\boldsymbol{\xi}_0)\le$$
$$\lambda\int_0^{+\infty}x\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x.$$

由$\int_0^{+\infty}x\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x=-\frac{1}{\sqrt{2\pi}}\int_0^{+\infty}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}\left(-\frac{x^2}{2}\right)=\frac{1}{\sqrt{2\pi}}$，可得

$$\lambda\left(\frac{1}{2}\right)^{\lambda-1}\frac{1}{\sqrt{2\pi}}\le\mathbb{E}(\eta_1|\boldsymbol{\xi}_0)\le\lambda\frac{1}{\sqrt{2\pi}}$$

假定算法从原点出发，此时$d(\boldsymbol{\xi}_0)=a$，根据定理1，立即得到

$$\frac{\sqrt{2\pi}(a-\varepsilon)}{\lambda}\le\mathbb{E}[\tau_\varepsilon|\boldsymbol{\xi}_0=(0,0)]\le\frac{\sqrt{2\pi}a}{\lambda\left(\frac{1}{2}\right)^{\lambda-1}}.$$

证毕.

需要指出的是本文的方法无需事先对η_t，$t=1,2,\cdots$作独立同分布的假设，只是在分析这个具体实例时恰好可以证明η_t，$t=1,2,\cdots$是独立同分布的，从而使得计算简化. 而文献[21]的模型事先必须作η_t，$t=1,2,\cdots$独立同分布的假设，所分析的算法与实例又恰好满足这一假设. 由此可见本文的方法比文献[21]的方法更具一般性.

5　结　论

由于进化算法的首达时间实质是一个随机停止时间，故引入停时来刻画是自然的想法. 本文引入停时理论，通过构造合适的距离函数，运用时齐马氏过程的性质分析了平均首达时间、算法的一步平均位移、初始种群与最优种群空间的距离这三者的关系，得到了估计平均首达时间的一般方法. 为展示这一方法的应用，我们选择分析更为复杂的(1+λ)EA和(1+λ)ES在两个问题实例上的计算时间，这与以往大部分的工作都是分析(1+1)EA、(1+1)ES不同，结果表明该方法适用于一般的进化算法.

通过分析比较(1+λ)EA和(1+1)EA在PEAK函数上的平均首达时间，我们发现增大子代种群的规模无助于降低(1+λ)EA求解PEAK函数的计算复杂度. 对于倾斜平面问题，文献[21]建立的更新过程模型要求对算法的一步位移作独立同分布的假设，这是很强的假设，在实际中能满足这一假设的情况很少，只是他们所分析的实例恰好满足这一假设. 本文所建立的理论模型不需作这样的假设，从而更具通用性. Level-reaching Estimation Technique是分析离散型进化算法计算时间的一个简单而又威力巨大的工具，得到广泛的应用，但原有的证明不够严谨，在本文中可以看出它只是我们所提出的理论模型的一个特例，从而得到了严格的证明.

最后要指出的是，本文所建立的理论模型不仅适用于进化算法的计算时间分析，经过改进推广后也适用于ACO、PSO等仿生优化算法的计算时间分析. 未来将要进一步研究：(1)改进本文的理论模型，将其推广到其他仿生算法的计算时间分析中；(2)分析进化算法在具有实际背景的组合优化问题上的计算时间；(3)分析种群规模大于1的进化算法，尤其是带交叉算子的遗传算法在一些优化问题实例上的计算时间. 这些都将是极具挑战性的工作.

参　考　文　献

[1] Oliveto P S, He J, Yao X. Time complexity of evolutionary algorithms for combinatorial optimization: A decade of results. International Journal of Automation and Computing, 2007, 4(3): 281-293

[2] Lehre P K, Yao X. Runtime analysis of the (1+1) EA on computing unique input output sequences. Information Sciences, 2014, 259(1): 510-531

[3] Zhang Y S, Hao Z F, Huang H, Lim A. Runtime analysis of (1+1) evolutionary algorithm for two combinatorial optimization instances. Journal of Information & Computational Science, 2011, 8(15): 3497-3506

[4] Droste S, Jansen T, Wegener I. On the analysis of the (1+1) evolutionary algorithm. Theoretical Computer Science, 2002, 276(1-2): 51-81

[5] Wegener I. Methods for the analysis of evolutionary algorithms on pseudo-boolean functions//Sarker R, Mohammadian M, Yao X eds. Evolutionary Optimization. Norwell: Kluwer Academic Publishers, 2002: 349-369

[6] Oliveto P S, He J, Yao X. Analysis of the (1+1) EA for finding approximate solutions to vertex cover problems. IEEE Transactions on Evolutionary Computation, 2009, 13(5): 1006-1029

[7] Borisovsky P A, Eremeev A V. Comparing evolutionary algorithms to the (1+1)EA. Theoretical Computer Science, 2008, 403(1): 33-41

[8] He J, Yao X. From an individual to a population: An analysis of the first hitting time of population-based evolutionary algorithms. IEEE Transactions on Evolutionary Computation, 2002, 6(5): 495-511

[9] He J, Yao X. Towards an analytic framework for analysing the computation time of evolutionary algorithms. Artificial Intelligence, 2003, 145(1-2): 59-97

[10] Jansen T, Jong K A D, Wegener I. On the choice of the offspring population size in evolutionary algorithms. Evolutionary Computation, 2005, 13(4): 413-440

[11] Yu Y, Zhou Z H. A new approach to estimating the expected first hitting time of evolutionary algorithms. Artificial Intelligence, 2008, 172(15): 1809-1832

[12] Huang Han, Hao Zhi-Feng, Wu Chun-Guo, Qin Yong. The convergence speed of ant colony optimization. Chinese Journal of Computers, 2007, 30(8): 1343-1353(in Chinese)
(黄翰，郝志峰，吴春国，秦勇．蚁群算法的收敛速度分析．计算机学报，2007，30(8)：1343-1353)

[13] Huang Han, Hao Zhi-Feng, Qin Yong. Time complexity of evolutionary programming. Journal of Computer Research and Development, 2008, 45(11): 1850-1857(in Chinese)
(黄翰，郝志峰，秦勇．进化规划算法的时间复杂度分析．计算机研究与发展，2008，45(11)：1850-1857)

[14] Jägersküpper J. Combining Markov-chain analysis and drift analysis: The (1+1) evolutionary algorithm on linear functions reloaded. Algorithmica, 2011, 59(3): 409-424

[15] He J, Yao X. Drift analysis and average time complexity of evolutionary algorithms. Artificial Intelligence, 2001, 127(1): 57-85

[16] He J, Yao X. A study of drift analysis for estimating computation time of evolutionary algorithms. Natural Computing: An International Journal, 2004, 3(1): 21-35

[17] Oliveto P S, Witt C. Simplified drift analysis for proving lower bounds in evolutionary computation. Algorithmica, 2011, 59(3): 369-386

[18] Chen T, He J, Sun G, et al. A new approach for analyzing average time complexity of population-based evolutionary algorithms on unimodal problems. IEEE Transactions on System, Man, and Cybernetics—Part B: Cybernetics, 2009, 39(5): 1092-1106

[19] Jägersküpper J. Algorithmic analysis of a basic evolutionary algorithm for continuous optimization. Theoretical Computer Science, 2007, 379(3): 329-347

[20] Jägersküpper J. How the (1+1)ES using isotropic mutations minimizes positive definite quadratic forms. Theoretical Computer Science, 2006, 361(1): 38-56

[21] Agapie A, Agapie M, Baganu G. Evolutionary algorithms for continuous-space optimisation. International Journal of Systems Science, 2013, 44(3): 502-512

[22] Zhou Y R. Runtime analysis of ant colony optimization algorithm for TSP instances. IEEE Transactions on Evolutionary Computation, 2009, 13(5): 1083-1092

[23] Wang Yu-Ping. Theory and Method of Evolutionary Computation. Beijing: Science Press, 2011(in Chinese)
(王宇平．进化计算的理论和方法．北京：科学出版社，2011)

[24] Huang Han, Lin Zhi-Yong, Hao Zhi-Feng, et al. Convergence analysis and comparison of evolutionary algorithms based on relation model. Chinese Journal of Computers, 2011, 34(5): 801-811(in Chinese)
(黄翰，林智勇，郝志峰等．基于关系模型的进化算法收敛性分析与对比．计算机学报，2011，34(5)：801-811)

[25] Lin Yuan-Lie. Applied Stochastic Process. Beijing: Tsinghua University Press, 2002(in Chinese)
(林元烈．应用随机过程．北京：清华大学出版社，2002)

[26] Neumann F, Witt C. Bioinspired Computation in Combinatorial Optimization-Algorithms and Their Computational Complexity. New York: Springer-Verlag, 2010

ZHANG Yu-Shan, born in 1975, Ph.D., associate professor. His main research interests include theoretical foundation of evolutionary algorithms.

HAO Zhi-Feng, born in 1968, Ph.D., professor. His main research interests include design and analysis of algorithms, mathematical foundation of bio-inspired algorithms, combinatorial optimization and algebra.

HUANG Han, born in 1980, Ph.D., associate professor. His main research interests include theoretical foundation of evolutionary computation, optimization design and application of evolutionary computation.

LIN Zhi-Yong, born in 1977, Ph.D., associate professor. His main research interests include design and analysis of algorithms, machine learning and computational intelligence.

Background

The runtime analysis of evolutionary algorithms (EAs) is an important topic which currently attracts many scholars' attention, it has important implications for deeply revealing the operational mechanism of EAs, guiding the design, improvement and application of EAs. In contrast to fruitful research results of EAs in practical applications, theoretical results on the runtime analysis are still few. The inherent randomness of EAs make it hard to analyze the runtime. The current international researches mainly focus on seeking suitable mathematical tools, including the improvement of existing methods and introduction of novel mathematical tools to analyze the runtime of specific EAs on some instances. The main approaches at present include Markov model, drift analysis, Level-reaching Estimation Technique, renewal process, etc.

This article introduces the stopping time theory, regards the first hitting time of EAs as stopping time, proposes a general approach for the runtime analysis of EAs. This theoretical approach does not need to make too many assumptions like the renewal process, and has a wide range of application. Under this theoretical framework, Level-reaching Estimation Technique becomes a direct conclusion. The proposed approach can be extended to the runtime analysis of other bio-inspired algorithms, and it will be the next step direction of our research.

This research is supported by the National Natural Science Foundation of China (Nos. 61370177, 61202453), the Humanity and Social Science Youth Foundation of Ministry of Education of China (No. 14YJCZH216), the Natural Science Foundation of Guangdong Province in China (No. S2011040002890), the Foundation for Distinguished Young Talents in Higher Education of Guangdong (No. 2013LYM_0049), the Foundation for Scientific and Technological Innovation in Higher Education of Guangdong (No. 2013KJCX0117). Our research group conducts research around the mathematical foundations of bio-inspired algorithms in recent years, and has achieved some theoretical achievements on the convergence and time complexity analysis of some bio-inspired algorithms such as EAs, ACO, PSO. We have published several related research papers on authoritative journals at home and abroad such as Science China, Chinese Journal of Computers, Journal of Computer Research and Development, IEEE Transactions on System, Man, and Cybernetics—Part B: Cybernetics, Information Sciences.