关于漂移分析的注记

周育人，赖鑫生
ZHOU Yuren, LAI Xinsheng
华南理工大学 计算机科学与工程学院，广州 510640
School of Computer Science and Engineering, South China University of Technology, Guangzhou 510640, China

ZHOU Yuren, LAI Xinsheng. Note on drift analysis. Computer Engineering and Applications, 2012, 48（8）：21-23.

Abstract： In spite of many successful applications of evolutionary algorithms in various complicated optimization problems, the rigorous theoretical analyses of EAs are still in their infant period. The running time of an evolutionary algorithm for combination optimization problem is an important topic in theoretical study of EAs. This paper discusses the relation between drift analysis and evolutionary algorithms' time complexity. It gives a new proof of the drift theorem. It also uses drift theorem to estimate the expected runtime of global（1+1）for linear function. The result is a better understanding of the performance of evolutionary algorithms.

Key words： evolutionary algorithms; time complexity; drift analysis

摘　要：进化算法成功应用于求解各种复杂优化问题，其理论研究尚处于初级阶段。时间复杂性分析可以估计算法的平均运行时间，是进化算法理论研究中的重要方向和有力工具。讨论了漂移分析和进化算法时间复杂性的关系，利用吸收马尔科夫链给出漂移定理的一个新的证明；用一步平均漂移估计算法计算时间，得到了线性函数进化算法时间复杂度的一个一般性的结果。这些结果有助于更好地理解进化算法的工作原理和性能。

关键词：进化算法；时间复杂性；漂移分析

DOI：10.3778/j.issn.1002-8331.2012.08.006　　**文章编号：**1002-8331（2012）08-0021-03　　**文献标识码：**A　　**中图分类号：**TP18

1　引言

作为现代启发式算法（Heuristic algorithms）的一个重要分支，进化算法（Evolutionary Algorithms，EAs）广泛应用于现实世界的优化问题；随着进化算法研究的深入，其理论研究日益受到重视[1-2]。

使用二进制编码的遗传算法的早期理论研究工作之一是模式理论和建筑块假设[3-4]，它们注重算法结构研究，对算法的收敛性、时间复杂性则几乎没有涉及。另外一些理论结果是使用马氏链模型分析遗传算法[5]，马氏链模型具有直观、对算法解释能力强等特征；早期的马氏链分析没有考虑算法的适应值函数，没有利用遗传算法的选优机制，得到的收敛性结果本质上无异于随机抽样算法相应的结果；早期的马氏链模型也没有考虑算法的时间复杂性。

组合优化问题进化算法的计算时间，即随机算法到达最优点所需平均进化代数，是进化算法理论研究的一个重要课题。当今的计算机科学理论在计算复杂性理论上取得了丰富的研究成果，开发了许多强有力的工具估计算法计算时间。多项式时间可解问题和指数问题的分界是复杂性理论的主要论题，作为随机启发式算法的进化算法也离不开这个重要主题。

当前进化算法时间复杂性分析的主要工作有：Rudolph[6]研究了长路径问题（Long path problems）计算时间；Garnier[7]分析了只有变异没有选择的算法时间复杂度。Droste、Wegener等学者分析了简单的（1+1）-EA在不同函数类上的计算时间，如单峰函数（Unimodal functions）[8]，线性函数（Linear functions）[9]，二次多项式（Quadratic polynomials）[10]；He Jun和Yao Xin用漂移分析（Drift analysis）[11-12]和吸收马尔科夫链（Absorbing Markov chains）[13]模型分析进化算法的时间复杂性以及群体在进化算法中的作用[14]；Laumanns[15]研究了多目标进化算法的时间复杂性，Yuren Zhou和Jun He[16]等研究了进化算法求解约束优化问题的时间复杂度，分析了惩罚参数在时间复杂度分析中的作用；以及分析和比较了（1+1）EA、Randomwalk等不同启发式算法求解SAT问题的时间复杂度[17]。

本文讨论了进化算法的时间复杂性和漂移分析关系。介绍了用一步平均漂移估计算法计算时间，给出了漂移定理的一个新证明，得到了线性函数进化算法时间复杂度的一个一般性的结果。

2　进化算法和吸收马尔科夫链

2.1　（1+1）EA

考虑以下拟布尔（pseudo-boolean）函数优化问题（最大值）：

$$\text{Maximize } f: \{0,1\}^n \to \mathbf{R} \tag{1}$$

群体规模为1且只使用变异和选择操作的进化算法简记为（1+1）EA，其一般描述为：

算法1　（1+1）EA

（1）Initialization

Choose uniformly at random.

（2）Mutation

y:=mutate(x).

（3）Selection

If $f(y)>f(x)$, $x:=y$.

（4）Stopping Criterion

If the stopping criterion is not met，continue at line 2.

（1+1）EA通常使用具有局部和全局性质的两种变异，分别称为局部变异（Local Mutation，LM）和整体变异（Global

基金项目：国家自然科学基金（No.61165003，61170081，60873078）；广东省自然科学基金（No.9251064101000010）。
作者简介：周育人（1965—），男，教授，博导，主要研究方向：计算智能；赖鑫生（1972—），男，博士生。E-mail：yrzhou@scut.edu.cn
收稿日期：2011-06-28；**修回日期：**2011-11-18；**CNKI出版：**2012-01-05；http://www.cnki.net/kcms/detail/11.2127.TP.20120105.1706.033.html

Mutation，GM）。

Local Mutation（LM）：在个体 $x\in\{0,1\}^n$ 中随机选取某一位取反（将 x_i 变成 $1-x_i$）。

Global Mutation（GM）：对个体 $x\in\{0,1\}^n$ 中每一位以变异概率 $1/n$ 取反。

整体变异中的变异概率可以任取一个概率值，在理论分析和实际应用中通常取为 $1/n$ 。

（1+1）EA是一个简单有效的随机爬山进化算法，是进化算法时间复杂性分析的一个基本分析模型；其时间复杂性分析方法和结果有重要的理论意义，也对更一般、更复杂的进化算法的理论分析有指导作用。

2.2 吸收马尔科夫链

一般地，进化算法的群体序列是一个离散马尔科夫链，（1+1）EA群（个）体是一个以最优解集为吸收集的吸收马尔科夫链。下面介绍一些齐次吸收马尔科夫链的基本知识，其细节可在一般的“随机过程”文献（如文献[18]）中找到。

设 $(X_t;t=0,1,\cdots)$ 为有限离散状态空间 S 上的离散吸收马尔科夫链，T 为表示瞬时状态集，$H=S-T$ 为吸收状态集。规范转移概率矩阵（Canonical transition matrix）为：

$$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{R} & \boldsymbol{T}\end{pmatrix}$$

其中 $\boldsymbol{I}$ 为单位矩阵，$\boldsymbol{T}$ 为瞬时态集的转移矩阵，$\boldsymbol{R}$ 为瞬时态到吸收态的转移矩阵。在 $\boldsymbol{P}$ 的任意次幂中，分块矩阵 $\boldsymbol{I}$ 保持不变，这对应着吸收马尔科夫链的特征：一旦到达吸收状态集 H，则永不离开。

定义1 设 $(X_t;t=0,1,\cdots)$ 为一个吸收马尔科夫链，从状态到 $i\in S$ 到达吸收状态集 H 的首达时间为：

$$\tau_i=\min\{t:t\geq 0,\ X_t\in H|\ X_0=i\}$$

若右边为空集，则令 $\tau_i=\infty$ 。

当 $i\in H$ 时，显然有 $\tau_i=0$。首达时间 τ_i 是一个随机变量，感兴趣的是它的数学期望 $E[\tau_i]$ 。

定理1 设 $(X_t;t=0,1,\cdots)$ 为一个吸收马尔科夫链，τ_i 是瞬时状态 i 的首达时间，记 $m_i=E[\tau_i]$ ，$\boldsymbol{m}=[m_i]_{i\in T}$ ，则：

$$\boldsymbol{m}=(\boldsymbol{I}-\boldsymbol{T})^{-1}\boldsymbol{1}$$

其中 $\boldsymbol{1}$ 表示向量 $(1,1,\cdots,1)^{\mathrm{t}}$ 。

证明 见文献[18]。

定理1给出了由转移概率求吸收马尔科夫链平均首达时间的解析表达式，然而由于进化算法所决定的吸收马尔科夫链转移矩阵往往过于复杂，直接使用定理1估计进化算法的时间复杂性并非易事。下面由定理1导出漂移定理，以便能估计进化算法时间复杂性。

3 漂移分析、收敛速率和进化算法时间复杂性

进化算法的一步漂移为算法运行一代时，群体朝最优点逼近时所获得的前进量，它描述了进化算法的局部行为。Sasaki和Hajek[19]首次将漂移分析应用于最大匹配问题的模拟退火解法的时间复杂性分析；He Jun和Yao Xin[11-12]将其应用于进化算法时间复杂性分析。

如前所述，设 S 为优化问题的状态空间，H 为最优解集，ξ_t 为（1+1）EA的第 t 代群体（由于群体规模为1，ξ_t 实际上为一个个体）。为了衡量群体 ξ_t 到最优解集 H 的距离，需要引入一个 S 上的距离函数 $d(x)$，$d(x)$ 有多种不同定义方法，如 $d(x)=\min\{|f(x)-f(y)||y\in H\}$，或 $d(x)$ 为 x 到最优解集 H 的哈明距离等。$D(x)$ 应满足当 $x\in H$ 时，$d(x)=0$。

定义2 （1+1）EA的第 t 代群体 ξ_t 在 x 处的一步平均漂移定义为：

$$\Delta d(\xi_t)|_{\xi_t=x}=E[d(\xi_t)-d(\xi_{t+1})|\xi_t=x]=$$
$$d(x)-\sum_{y\in S}p(x,y)d(y)=d(x)-\sum_{y\in H}p(x,y)d(y)$$

其中 $p(x,y)$ 为进化算法从状态 x 到状态 y 的转移概率。

一步平均漂移可以用来估计算法找到最优解的平均计算时间。

定理2 （1）设 $d(x)$ 为状态空间 S 上的距离函数，若对于某个 $x_0\in S$ ，有 $d(x_0)\leq h_0(n)$ ；且 $\forall x\in S$ ，（1+1）EA的一步平均漂移 $\Delta d(\xi_t)|_{\xi_t=x}\geq h_1(n)$ ，则算法从初始状态 x_0 出发到达最优解的平均首达时间：

$$E[\tau_{x_0}]\leq\frac{h_0(n)}{h_1(n)}$$

（2）设 $d(x)$ 为状态空间 S 上的距离函数，若对于某个 $x_0\in S$ ，有 $d(x_0)\geq h_0(n)$ ；且 $\forall x\in S$ ，（1+1）EA的一步平均漂移 $\Delta d(\xi_t)|_{\xi_t=x}\leq h_1(n)$ ，则算法从初始状态 x_0 出发到达最优解的平均首达时间：

$$E[\tau_{x_0}]\geq\frac{h_0(n)}{h_1(n)}$$

证明 （1）$d(x)$、$\Delta d(\xi_t)$ 在 $S-H$ 上的值组成的向量分别记为 $\boldsymbol{d}=[d_i]_{i\in S-H}$ 和 $\boldsymbol{\Delta}=[\Delta_i]_{i\in S-H}$ ，则有：

$$\boldsymbol{\Delta}=(\boldsymbol{I}-\boldsymbol{T})\boldsymbol{d}$$

设 $(\boldsymbol{I}-\boldsymbol{T})^{-1}=(a_{ij})$ ，由 $\boldsymbol{d}=(\boldsymbol{I}-\boldsymbol{T})^{-1}\boldsymbol{\Delta}$ 得：

$$d_i=\sum_j a_{ij}\Delta_j$$

由于 $\Delta_j\geq h_1(n)$ ，有：

$$d_i\geq h_1(n)\sum_j a_{ij}$$

即：

$$\sum_j a_{ij}\leq\frac{d_i}{h_1(n)}$$

又由定理1，$\boldsymbol{m}=(\boldsymbol{I}-\boldsymbol{T})^{-1}\boldsymbol{1}$ ，得：

$$m_i=\sum_j a_{ij}\leq\frac{d_i}{h_1(n)}$$

所以

$$E[\tau|\xi_0=X]\leq\frac{h_0(n)}{h_1(n)}$$

（2）类似于（1），略。

定理2即为文献[11]中的定理3和定理4，是漂移分析估计进化算法计算时间的主要结果和工具，本文给出了一个新的简洁的证明，这里没有使用鞅、停时等随机过程中的一些“高深”概念。

应用定理2来估计进化算法时间复杂性的关键在于适当地选择距离函数 $d(x)$，以下将使用定理2估计进化算法求解常见的线性函数的平均计算时间。

定义3 设 $x=(x_1,x_2,\cdots,x_n)\in\{0,1\}^n$ ，定义线性函数为：

$$f(x)=a_1x_1+a_2x_2+\cdots+a_nx_n\ ,\ a_k>0\ ,k=1,2,\cdots,n$$

对于线性函数，Droste[10]和He[12]分别证明了在整体变异概率为 $1/n$ 、$1/(2n)$ 时，（1+1）EA时间复杂性为 $O(n\log n)$。下面

的结果表明，在整体变异概率为 $r/n(0<r<1)$ 时，结论也成立。

定理3 使用整体变异概率 $r/n(0<r<1)$ 的（1+1）EA求解线性函数的平均计算时间为（$n\log n$）。

证明 利用定理2证之。

设 $x=(x_1,x_2,\cdots,x_n)\in S=\{0,1\}^n$，记 $|x|=x_1+x_2+\cdots+x_n$。将搜索空间 S 划分成 $n+1$ 个子集 $S_m(0\le m\le n)$，其中 $S_m:=\{x|x\in S, n-|x|=m\}$。在 S 上定义距离函数如下：

$$d(x)=\begin{cases}0, x\in S_0\\ n(1+\cdots+\frac{1}{m}), x\in S_m, 0<m\le n\end{cases}$$

记

$$d_m=\begin{cases}0, m=0\\ n(1+\cdots+\frac{1}{m}), 0<m\le n\end{cases}$$

设 $i,j\in S$，$p(i,j)$ 为进化算法从状态 i 到 j 的转移概率，则当 $i\in S_m$ 时，其一步平均漂移为：

$$\Delta d(\xi_t)|_{\xi_t=i}=\sum_{k=0}^{n}\sum_{j\in S_k}(d_m-d_k)p(i,j)=$$

$$\sum_{k=1}^{m}(d_m-d_{m-k})\sum_{j\in S_{m-k}}p(i,j)+\sum_{v=1}^{n-m}(d_m-d_{m+v})\sum_{j\in S_{m+v}}p(i,j)$$

对于 $i\in S_m$，当 i 中 $k(k=1,2,\cdots,m)$ 个0发生变异，且其余位保持不变时，所产生的个体 $j\in S_{m-k}$ 满足 $f(j)>f(i)$，所以：

$$\sum_{j\in S_{m-k}}p(i,j)\ge\binom{m}{k}\left(\frac{r}{n}\right)^k\left(1-\frac{r}{n}\right)^{n-k}$$

而对于 $i\in S_m$，仅当 i 中 $k(k=1,2,\cdots,m)$ 个0发生变异，同时 $k+v(v=1,2,\cdots,n-m)$ 个1发生变异，且其余位保持不变时，所产生的个体 $j\in S_{m+v}$ 才有可能满足 $f(j)>f(i)$，故：

$$\sum_{j\in S_{m+v}}p(i,j)\le\sum_{k=1}^{m}\binom{m}{k}\binom{n-m}{k+v}\left(\frac{r}{n}\right)^{2k+v}\left(1-\frac{r}{n}\right)^{n-2k-v}$$

从而

$$\begin{aligned}\Delta d(\xi_t)|_{\xi_t=i}=&\sum_{k=1}^{m}(d_m-d_{m-k})\sum_{j\in S_{m-k}}p(i,j)+\\&\sum_{v=1}^{n-m}(d_m-d_{m+v})\sum_{j\in S_{m+v}}p(i,j)\ge\\&\sum_{k=1}^{m}\binom{m}{k}\left(\frac{r}{n}\right)^k\left(1-\frac{r}{n}\right)^{n-k}((d_m-d_{m-k})+\\&\sum_{v=1}^{n-m-k}(d_m-d_{m+v})\binom{n-m}{k+v}\left(\frac{r}{n-r}\right)^{k+v})\end{aligned}$$

当 $k\ge2$ 时，

$$(d_m-d_{m-k})+\sum_{v=1}^{n-m-k}(d_m-d_{m+v})\binom{n-m}{k+v}\left(\frac{r}{n-r}\right)^{k+v}\ge$$

$$n((\frac{1}{m}+\cdots+\frac{1}{m-k+1})-\sum_{v=1}^{+\infty}(\frac{1}{m+1}+\cdots+\frac{1}{m+v})\frac{1}{(k+v)!}r^{k+v})\ge$$

$$n\frac{1}{m}(1-\sum_{v=1}^{+\infty}\frac{v}{(v+1)!}r^{v+1})=n\frac{1}{m}(1-r)\mathrm{e}^r>0$$

这里使用了等式 $\sum_{v=1}^{+\infty}\frac{v}{(v+1)!}r^{v+1}=1+(r-1)\mathrm{e}^r$。

所以

$$\Delta d(\xi_t)|_{\xi_t=i}\ge$$

$$\binom{m}{1}\left(\frac{r}{n}\right)\left(1-\frac{r}{n}\right)^{n-1}(\frac{n}{m}+\sum_{v=1}^{n-m-1}(d_m-d_{m+v})\binom{n-m}{v+1}\left(\frac{r}{n-r}\right)^{v+1})$$

又因为

$$\sum_{v=1}^{n-m-1}(d_m-d_{m+v})\binom{n-m}{v+1}\left(\frac{r}{n-r}\right)^{v+1}\ge-\frac{n}{m}\sum_{v=1}^{+\infty}\frac{v}{(v+1)!}r^{v+1}=$$

$$-\frac{n}{m}(1+(r-1))\mathrm{e}^r)$$

所以

$$\Delta d(\xi_t)|_{\xi_t=i}\ge r(1-r)\mathrm{e}^{r-1}$$

考虑到距离函数 $d(x)=O(n\log n)$，由定理2有：

$$E[\tau|\xi_0=X]=O(n\log n)$$

4 结语

当前进化算法的理论研究中，进化算法时间复杂性研究尚处于初级阶段，时间复杂性分析的有力工具和一般分析框架有待挖掘和完善。本文仅讨论了（1+1）EA求解线性函数等人工问题的计算时间，进化算法在真实世界复杂问题中的应用，如旅行商、约束优化问题、多目标优化问题等，算法的时间复杂性值得进一步研究。有理由相信，进化算法时间复杂性理论的研究将丰富和发展现有的时间复杂性理论。

参考文献：

[1] Eiben A E，Rudolph G.Theory of evolutionary algorithms：a bird's eye view[J].Theoretical Computer Science，1999，229（1/2）：3-9.

[2] Beyer H G，Schwefel H P，Wegener I.How to analyse evolutionary algorithms[J].Theoretical Computer Science，2002，287（1）：101-130.

[3] Goldberg D E.Genetic algorithms in search，optimization and machine learning[M].Boston，MA，USA：Addison Wesley，1989.

[4] Holland J H.Adaptation in natural and artificial systems[M]. Ann Arbor：University of Michigan Press，1975.

[5] Nix A E，Vose M D.Modeling genetic algorithms with Markov chains[J].Annals of Mathematics and Artificial Intelligence，1992，5（1）：79-88.

[6] Rudolph G.How mutation and selection solve long path problems in polynomial expected time[J].Evolutionary Computation，1996，4（2）：194-205.

[7] Garnier J，Kallel L，Schoenauer M.Rigorous hitting times for binary mutations[J].Evolutionary Computation，1999，7（2）：167-203.

[8] Droste S，Jansen T，Wegener I.On the optimization of unimodal functions with the（1+1）evolutionary algorithm[C]//LNCS 1498：Proc of the 5th Conf on Parallel Problem Solving from Nature（PPSN Ⅴ），1998：13-22.

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

[10] Wegener I，Witt C.On the analysis of a simple evolutionary algorithm on quadratic pseudo-Boolean function[J].Journal of Discrete Algorithm，2005（3）：61-78.

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

[12] He J，Yao X.A study of drift analysis for estimating computation time of evolutionary algorithms[J].Natural Computation，2004，3：21-35.

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

（下转27页）