

一种遗传进化规划

刘芳 刘民 吴澄

(清华大学自动化系 北京 100084)

摘要 本文提出一种遗传进化规划,该方法结合了遗传算法和进化规划两种算法的优点,在进化过程中遗传算法的交换率、变异率和进化规划的变异规则均根据种群的进化信息而自适应变化。该方法不仅能够加快算法的收敛速度,而且能够有效地保持种群的多样性。用该方法求解混合非线性整数规划问题,计算机仿真实验结果表明是非常有效的。

关键词 遗传算法,进化规划,进化性,混合非线性整数规划问题,熵

A Genetic Evolutionary Programming

LIU Fang LIU Min WU Cheng

(Department of Automation, Tsinghua University, Beijing 100084)

Abstract A genetic evolutionary programming is proposed in the paper, which combines the advantages of genetic algorithm and evolutionary programming. In the evolving process, the exchange rate and mutate rate of genetic algorithm and the mutate rules of evolutionary programming are changed self-adaptively according to evolution information of the population. It can not only keep the population diversity but also has quicker convergence speed. It is applied to integer programming. Computer simulation results show its validity.

Keywords Genetic algorithm, Evolutionary programming, Capability of evolution, Mixed integer non-linear programming problems, Entropy

1 前言

某些工程优化问题可以归结为混合非线性整数规划。虽然数字计算机对于一个含有数以 10 万个变量的线性规划可以在合理的时间内求得其最优解,但对于混合非线性整数规划(MINLP)却没有那么幸运。实际上,我们没有一个很好的求解整数规划和非线性规划的通用方法,混合非线性整数规划被认为是一类困难问题。近年来,进化算法用于求解混合非线性整数规划的研究引起人们的广泛关注,其中 GA(遗传算法)、ES(进化策略)、SA(模拟退火)等方法被成功用于求解此类问题。最近的研究成果认为 ES 求解 MINLP 问题的效果最好^[1]。

进化规划(EP)是进化算法的一种,与 GA 和 ES 相比,其遗传操作较简单且不需要编码(直接采用十进制编码)。在搜索性能上,EP 和 ES 的局部搜索性能相对较好,而 GA 的全局搜索性能相对较好。另外,ES 和 EP 都较 GA 花费更多的进化时间。若能将擅长大范围搜索的 GA 与擅长局部搜索的 EP 结合起来,可起到取长补短、相辅相成的作用。基于这种思路,我们提出一种新的进化算法(GAEP),该方法能够提高 EP 的全局收敛速度,并且有效地防止 EP 的早熟收敛。仿真试验表明,该方法用于求解混合非线性整数规划问题是有效的。

本文研究如下标准形式的 MINLP 问题:

$$\begin{aligned} & \min f(x) \\ & \text{s. t} \\ & h(x)=0 \end{aligned}$$

$$\begin{aligned} & g(x) \leq 0 \\ & x_j, \text{int } g_j, j \in I \\ & x \in X = \{x | x \in R^n, x^l \leq x \leq x^u\} \\ & \text{where} \\ & h \in R^m, g \in R^p \end{aligned}$$

2 GAEP 算法

2.1 遗传算法和进化规划的比较

遗传算法(以下简称为 GA)是进化算法中最流行的和宏观意义下的一种仿生算法,它的应用研究比理论研究更为丰富,已渗透到许多学科和领域。遗传算法模仿的机制是生命和智能的产生与进化过程,选择基于每个个体对外部环境的适应性,即遵循达尔文“优胜劣汰,适者生存”的自然法则,同时考虑了进化过程中的随机扰动情况。遗传算法起源于上世纪 60 年代对自然和人工适应系统的研究,最早由美国密执安大学的 Holland 教授提出。70 年代,De Jong 基于遗传和进化思想在计算机上进行了大量的纯数值函数的计算^[2]。在一系列研究工作的基础上,80 年代 Goldberg 进行了归纳总结,形成了遗传算法的基本框架。目前,遗传算法的研究依然非常活跃,已经形成了一套比较完整和成熟的理论,反映出它在优化计算中的重要地位^[3,4]。

1962 年,美国的 L. J. Fogel 首先提出进化规划(Evolutionary Programming,简称 EP),当时未得到足够的重视。30 多年后,其子 D. B. Fogel 改善了这种方法,从而使进化规划作为进化算法的一个分支得到广泛应用。与遗传算法相比较,进化规划强调的是父代与其子代在个体层、物种层的行为联

系,而遗传算法强调的是进化操作。相对而言,进化规划的应用不如遗传算法、遗传规划及进化策略广泛,相应的理论研究也比较薄弱^[4],但是关于 EP 的研究正在日益引起人们的重视。EP 这种进化优化方法的突出特点是^[4]:第一,能够渐近收敛到问题的全局最优解;第二,进化曲线变化比较平缓,进化求优时间一般很长。因而,基于实数编码的进化规划更适合于高精度的局部搜索。

2.2 遗传进化规划算法(GAEP)

遗传算法中存在一个重要问题是成熟前收敛(premature convergence),它使遗传算法搜索不到真正的最优值。这时群体中的所有染色体都相同,以致于基因交换只能产生和父代相同的子代。任何进一步的改善只能依靠基因突变,而突变又是遗传算法中的次要进化算子,(一般来说突变概率为 0.001 左右),因而收敛速度变得极慢^[5]。若将以突变算子为主要进化算子的进化规划与遗传算法相结合,则可以改善遗传算法的早熟收敛问题。另一方面,从对进化规划算法的进一步分析中不难发现:变异算子是在一定的概率分布条件下,随机地、没有指导地对个体进行修正。因此,它在为群体中的个体提供了进化机会的同时,也不可避免地产生了退化的可能。在某些情况下,特别是对一些难度较大的问题求解,这种退化现象会更加明显。另外,每一个待求的实际问题都会有自身的一些基本的、显而易见的特征信息或知识,而进化规划算法的变异算子在求解问题时,可变的灵活程度较小。算法的这种特点无疑对其通用性是有益的,但却忽视了问题的特征信息对求解问题所起到的辅助作用,特别是求解一些复杂问题时,这种“忽视”所带来的损失往往就更加明显了。我们提出利用遗传算法“获得”的局部特征信息以一定的强度干预进化规划的搜索进程、抑制或避免求解过程中的一些重复和无效的工作,以克服原进化规划算法中变异操作的盲目性。

2.2.1 进化算法的进化性的度量

进化性是评价进化算法性能的重要指标,它体现的是如何从算法内在机制上保证得到问题的满意解。进化性是指在进化过程中,种群产生的后代比其父代具有更好的适应环境的能力,它是针对进化算法而言的,由种群产生的下一代个体及其适应值的分布来衡量。在进化过程中,父代种群的性能与算子的作用决定了子代种群的适应值分布。如果能够确定子代种群适应值的概率分布,那么相应的适应值分布的数字特征,如均值、方差等,便可立刻获得,就可对算法的进化性作出定量评判。实质上,进化过程中,种群的多样性是进化算法进化性好坏的表现。我们研究种群的进化性本质上也是为了保持种群的多样性。我们参考文^[6],采用种群的方差和熵共同度量种群在基因空间的多样性。其中,第 t 代种群的熵定义如下:

设种群第 t 代种群按个体适应值的不同分为 Q 个子集: $S_1(t), S_2(t), \dots, S_Q(t)$, 各个子集所包含的个体数目记为 $|S_1(t)|, |S_2(t)|, \dots, |S_Q(t)|$, 且对任意 $p, q \in \{1, 2, \dots, Q\}$, 有

$$S_p(t) \cap S_q(t) = \emptyset, \bigcup_{q=1}^Q S_q(t) = A(t)$$

其中, $A(t)$ 为第 t 代种群的集合。则定义第 t 代种群的熵如下:

$$E(t) = - \sum_{j=1}^Q p_j(t) \log(p_j(t))$$

其中, $p_j(t) = \frac{|S_j(t)|}{N}$; N 为种群规模。

第 t 代种群的熵可以看作是第 t 代种群中各个子集的分

布概率的加权统计平均值,是不确定性的一种度量,它度量的是第 t 代种群中个体分布的“混乱程度”。

第 t 代种群的方差定义如下:

令第 t 代种群中的第 i 个个体适应值为 $f_i(t)$, 定义第 t 代种群的平均适应值 $\bar{f}(t)$ 为:

$$\bar{f}(t) = \frac{1}{N} \sum_{i=1}^N f_i(t)$$

则第 t 代种群的方差定义为:

$$D(t) = \frac{1}{N} \sum_{i=1}^N (f_i(t) - \bar{f}(t))^2$$

其中, N 为种群规模。

第 t 代种群的方差也是种群不确定性的一种度量,但它度量的是第 t 代种群中个体分布的“分散程度”。种群的方差和熵从不同的侧面共同反映了种群在基因空间的多样性。

2.2.2 GAEP 的自适应交换率、自适应变异率和自适应进化步长

在 GAEP 算法中, GA 和 EP 的任务不同,因而对算法的要求也不同。GA 注重搜索未知区域,要求处理的信息多,处理速度快而对精度要求不高; EP 注重搜索有局部最优点的区域,要求搜索速度快,一般来说其搜索范围较小。但这两种算法何时切换? 如何切换是我们算法设计的关键。

设 $E(t)$ 和 $\sigma(t)$ 为第 t 代种群的多样性度量指标(熵和方差)。当 $E(t)$ 和 $\sigma(t)$ 都很大时,表明第 t 代种群中的个体不但十分分散而且各不相同(“较为混乱”),此时种群的进化性最大,适合使用交换算子进行大范围搜索;当 $E(t)$ 很小而 $\sigma(t)$ 很大时,则表明第 t 代种群中的个体虽然有较分散的分布,但有很多几乎相互重叠的个体(“不够混乱”),此时种群的进化性不够好,宜采用较大变异率的 GA 进行搜索;当 $E(t)$ 很大而 $\sigma(t)$ 很小时,则表明第 t 代种群中的个体分布较为集中,但个体之间各不相同(“较为混乱”),此时种群的进化性较好,宜采用 EP 进行局部细搜;当 $E(t)$ 和 $\sigma(t)$ 都很小时,则表明第 t 代种群中的个体分布很集中,而且个体之间几乎相互重叠(“不够混乱”),此时的种群有收敛趋势,若经判断,尚未达到收敛条件,则宜采用较大变异的 EP 跳出局部最优点。在上述分析的基础上,我们提出新的 GAEP 算法如下:

if $E(t) \geq a$ & $\sigma(t) \geq b$

then $P_c = P_a - \text{generation}(t) * c, P_m = 1 - P_c$, 启用 GA。

else if $E(t) < a$ & $\sigma(t) \geq b$

then $P_c = P_m - c * \text{generation}(t), P_m = 1 - P_c$, 启用 GA。

else if $E(t) \geq a$ & $\sigma(t) < b$

then $r = r_1$, 启用 EP1。

else if $E(t) < a$ & $\sigma(t) < b$

else if 不满足算法中止条件, then $r = r_n$, 启用 EP2。

else 算法中止, 输出结果。

其中, $E(t)$ 表示第 t 代种群的熵; $\sigma(t)$ 表示第 t 代种群的方差; a, b, c 为可以调节的算法参数; P_c 表示遗传算法的交换率; P_m 表示遗传算法的变异率; P_a 表示较高的遗传算法交换率; P_m 表示较低的遗传算法交换率; P_m 表示较低的遗传算法变异率; P_m 表示较高的遗传算法变异率; $\text{generation}(t)$ 表示第 t 代的进化代数; r_1 表示 EP 的进化步长的系数; r_n 表示较大的 EP 的进化步长的系数; EP1 表示使用第一类变异规则的 EP, 参见公式(1); EP2 表示使用第二类变异规则的 EP, 参见公式(2)。

GAEP的工作流程类似于其他进化算法,同样经历产生初始群体-变异-计算个体适应度-选择-组成新群体,然后反复迭代,一代一代地进化,直至达到最优解。我们采用 q 竞争选择策略^[4]。

2.2.3 GAEP的进化算子设计

GA的进化算子。常规GA的交换算子和变异算子都是基于二进制编码的,因而其算法精度和性能受到二进制编码长度的限制。我们基于实数编码,设计GA的交换算子如下:

$$x_i(t+1) = x_i(t) + (x_j(t) - x_i(t)) * \eta$$

$$x_j(t+1) = x_j(t) + (x_i(t) - x_j(t)) * (1 - \eta)$$

其中, $x_i(t)$ 表示第 t 代种群中的第 i 个个体; $x_j(t)$ 第 t 代种群中的第 j 个个体; η 表示 $[0,1]$ 之间的一个随机数。

设计GA的变异算子如下:

$$x_i(t+1) = x_i(t) + (\eta - 0.5) * k$$

其中, $x_i(t)$ 表示第 t 代种群中的第 i 个个体; η 表示 $[0,1]$ 之间的一个随机数; k 为一调节参数。

EP的进化算子。EP的第一类变异规则要求具有快速收敛的特性。基于这一思路,设计EP的第一类变异规则如下:

$$x_i(t+1) = x_i(t) + (x_i(t) - x_{best}(t)) * r_1 * N(0,1) \quad (1)$$

其中, $x_i(t)$ 表示第 t 代种群中的第 i 个个体; $x_{best}(t)$ 表示第 t 代的整个种群中的最优个体; r_1 表示EP的进化步长的系数; $N(0,1)$ 表示一个服从正态分布的随机数。

EP的第二类变异规则要求具有快速跳出局部最优的特性。基于这一思路,设计EP的第二类变异规则如下:

$$x_i(t+1) = x_i(t) + k_i * N(0,1) * r_k \quad (2)$$

其中, $x_i(t)$ 表示第 t 代种群中的第 i 个个体; k_i 表示一个可以调节的参数,它与期望的EP的搜索范围有关; r_k 表示较大的EP的进化步长的系数; $N(0,1)$ 表示一个服从正态分布的随机数。

3 仿真实验

选用文[1]中的3个例子,来测试本文算法的性能。

1) $\min f(x,y) = 2x + y$

s. t $1.25 - x^2 - y \leq 0 \quad x + y \leq 1.6 \quad 0 \leq x \leq 1.6 \quad y \in \{0,1\}$

2) $\min f(x_1, x_2, x_3, y_1, y_2, y_3, y_4)$
 $= (y_1 - 1)^2 + (y_2 - 1)^2 + (y_3 - 1)^2 - \ln(y_4 + 1) + (x_1 - 1)^2 + (x_2 - 2)^2 + (x_3 - 3)^2$

s. t $y_1 + y_2 + y_3 + x_1 + x_2 + x_3 \leq 5$

$y_1^2 + x_1^2 + x_2^2 + x_3^2 \leq 5.5$

$y_1 + x_1 \leq 1.2$

$y_2 + x_2 \leq 1.8$

$y_3 + x_3 \leq 2.5$

$y_4 + x_1 \leq 1.2$

$y_2^2 + x_2^2 \leq 1.64$

$y_3^2 + x_3^2 \leq 4.25$

$y_2^2 + x_3^2 \leq 4.64$

$x_1, x_2, x_3 \geq 0$

$y_1, y_2, y_3, y_4 \in \{0,1\}$

3) $\max f(x_1, x_2, x_3, y_1, y_2)$

$= -5.357854x_1^2 - 0.835689y_1x_3 - 37.29329y_1 + 40792.141$

s. t $a_1 + a_2y_2x_3 + a_3y_1x_2 - a_4x_1x_3 \leq 92$

$a_5 + a_6y_2x_3 + a_7y_1y_2 + a_8x_1^2 - 90 \leq 20$

$a_9 + a_{10}x_1x_3 + a_{11}y_1x_1 + a_{12}x_1x_2 - 20 \leq 5$

$27 \leq x_1, x_2, x_3 \leq 45$

$y_1 \in \{78, \dots, 102\}, \text{integer}$

$y_2 \in \{33, \dots, 45\}, \text{integer}$

表1 问题(3)中的参数

a_1	85.334407	a_2	0.0056858	a_3	0.0006262	a_4	0.0022053
a_5	80.51249	a_6	0.0071317	a_7	0.0029955	a_8	0.0021813
a_9	9.300961	a_{10}	0.0047026	a_{11}	0.0012547	a_{12}	0.0019085

试验条件:求解问题(1)、(2)和问题(3)时,取种群规模均为50。用本文算法求解问题(1)的最优进化曲线如图1所示,求得的全局最优解为: $(x,y,f) = (0.5,1;2)$ 。用本文算法求解问题(2)的最优进化曲线如图2所示,求得的全局最优解为:

$(x_1, x_2, x_3, y_1, y_2, y_3, y_4; f) = (0.2, 1.280624, 1.954483, 3, 1, 0, 0, 1; 3.557463)$ 。用本文算法可求得问题(3)的全局最优解为:

$(x_1, x_2, x_3, y_1, y_2; f) = (27, 35.8175, 27, 78, 36, 3.22174 * 10^4)$,最优进化曲线图略。

文[1]中研究了GA和ES的不同参数(包括算法参数和约束问题处理方法)的大量实验,为节省篇幅,本文选择文[1]算法性能最好的 $ES(\mu+\lambda)$ 与本文算法做比较。重复试验60次,本文方法的收敛性能与文[1]中方法的收敛性能的比较见表2。

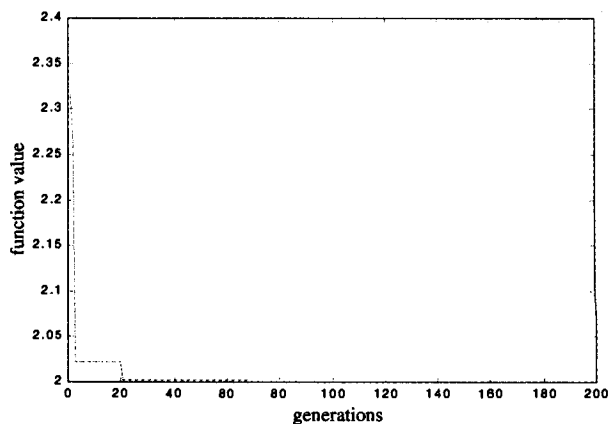


图1 问题(1)的最优进化曲线

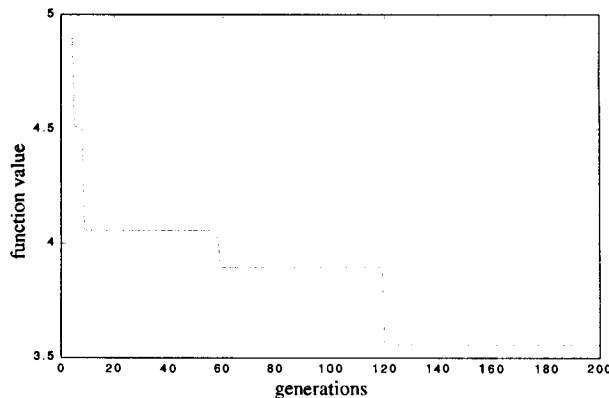


图2 问题(2)的最优进化曲线

根据以上分析可知:QoS-DRMR 采用了 MOIP 原则进行路由选择,发现,建立了保证最小带宽要求并取得整体优化延时和包丢失率条件的路由方案。保证了数据传输的带宽要求,同时也减少了 DRMR 路由的数目,从而减少了中间节

点转发数据包的数量,有效地减轻了网络负载,因而具有较高的数据传输率。因为 MOIP 原则是综合考虑了延迟和丢包率两个因素,所以能够获得较好的多目标整体优化效果。

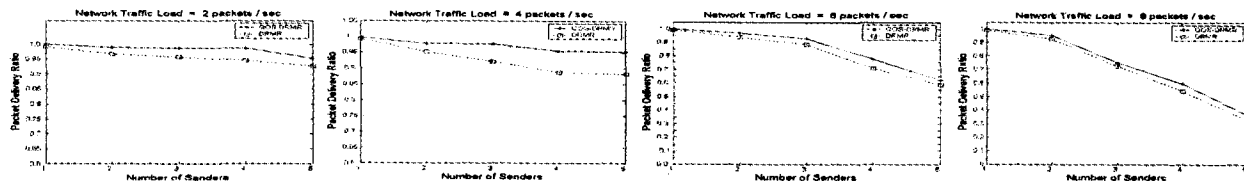


图 1 在不同送速率下源个数对两种协议性能的影响

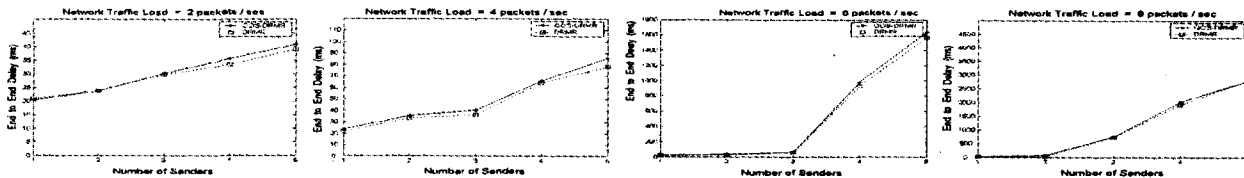


图 2 在不同发送速率下源个数对两种协议的端到端延迟的影响

结束语 本文根据 Ad Hoc 网络中组播对 QoS 保证的实际需求,考虑多目标约束条件,通过多目标整数优化模型来寻找满足最小带宽要求且具有优化的较小时延和较小包丢失率的组播路由,本文对 DRMR 协议中路由建立的过程进行了改造,形成了扩展的 DRMR 组播协议(QoS-DRMR),利用 NS2 仿真软件对 QoS-DRMR 协议进行仿真,结果表明:QoS-DRMR 协议较好地实现了在多目标约束条件下的组播服务。

进一步工作可以考虑支持多源节点对 QoS 要求不同情况下的路由建立机制以及通过交互机制进一步寻找更小时延和更小包丢失率的路由。

参 考 文 献

- 1 Agrawal D P, Zeng Qing-An. Introduction to Wireless and Mobile Systems [M]. Brooks/Cole, 2003
- 2 Royer E M, Perkins C E. Multicast Operation of the Ad-hoc On-Demand Distance Vector Routing Protocol [C]. In: Proc. of the Annual Intl. Conf. on Mobile Computing and Networking, MOBI-COM, 1999. 207~218
- 3 Wu C W, Tay Y C. AMRIS: A Multicast Protocol for Ad hoc Wireless Networks [C]. In: Proc. of IEEE MILCOM'99, Atlantic City, Nov. 1999. 25~29

- 4 Lee Sung-Ju, Gerla M, Chiang Ching-Chuan. On-Demand Multicast Routing Protocol in Multihop Wireless Mobile Networks [J]. Mobile Networks and Applications, 2002(7): 441~453
- 5 Garcia-Luna-Aceves J J, Madruga E L. The Core-Assisted Mesh Protocol [J]. IEEE Journal on Selected Areas in Communications, 1999, 17(8): 1380~1394
- 6 Zhou Yuan, Li Guang-Sheng, Zhan Yong-Zhao, et al. DRMR: Dynamic Ring Based Multicast Routing Protocol for Ad hoc Networks [J]. Journal of Computer Science and Technology, 2004, 19(6): 909~919
- 7 An Beongku, Papavassiliou S. A Mobility-Based Hybrid Multicast Routing in Mobile Ad-hoc Wireless Networks [C]. In: Proc. of MILCOM 2001. 316~320
- 8 Jetcheva J G, Johnson D B. Adaptive Demand-Driven Multicast Routing in Multi Hop Wireless Ad Hoc Networks [C]. In: Proc. of Mobihoc 2001, October Long Beach, CA, USA, 2001. 33~44
- 9 Fandel G, Gal T. Multiple Criteria Decision Making-Theory and Application [M]. New York: Springer-Verlag, 1980
- 10 Mi Zhi-chao, Zheng Shao-ren, Wang Ze-yan, et al. An Interactive QoS Routing Algorithm for Ad Hoc Networks [J]. Journal of Wuhan University (Natural Science Edition), 2002, 48(1): 051~054
- 11 Fall K, Varadhan K. NS notes and documentation. The VINT Project, UC Berkeley, LBL, USC/ISI and Xerox PARC, Nov. 1997. Available at: http://www.isi.edu/nsnam/ns/

(上接第 26 页)

表 2 本文算法与文[1]算法的性能比较

问题编号	本文算法的收敛性能	本文算法收敛的进化代数	ES($\mu+\lambda$)的收敛性能	ES($\mu+\lambda$)收敛的进化代数
问题 1	100%	200	100%	1518
问题 2	100%	1200	100%	6710
问题 3	100%	900	100%	2536

结论 本文提出一种遗传进化规划算法,实验表明该方法的自适应性能对于整数规划的困难问题有较好的搜索能力,并证明该方法优于现有的 GA、ES、SA 等方法。如何有效地应用遗传进化规划解决其它问题以及从理论上更深刻地研

究遗传规划的性能,将是我们继续研究的问题。

参 考 文 献

- 1 Costa L, Oliveira P. Evolutionary algorithms approach to the solution of mixed integer non-linear programming problems. Computers and chemical Engineering 2001, 25: 257~266
- 2 De Jong K. An analysis of the behavior of a class of genetic adaptive systems. [Ph D dis]. University of Michigan, 1975
- 3 Back T, Hammel U, Schwefel H-P. Evolutionary computation: Comments on the history and current state. IEEE Trans Evol Comput, 1997, 1(1): 3~17
- 4 云庆夏. 进化计算. 北京: 冶金工业出版社, 2001
- 5 李人厚. 智能控制理论和方法. 西安: 西安电子科技大学出版社, 1999
- 6 李孝安. 西北工业大学博士论文, 2000