

# 基于模拟退火的并行基因表达式编程算法研究

蒋思伟, 蔡之华, 曾 丹, 李 曲, 程远方

(中国地质大学计算机学院, 湖北武汉 430074)

**摘 要:** 本文提出一种基于 GEP 的混合同行算法, 将模拟退火机制与遗传机制结合, 以提高算法跳出局部最优的能力; 同时引入多种群策略和迁移算子, 实现了运行于一定规模处理器内的粗粒度可扩展并行计算. 对序列推理、函数发现和太阳黑子预测三组数据测试, 与用基本 GEP 算法的结果进行对比, 实验表明改进的算法具有更强的寻优能力和更高的稳定性.

**关键词:** 基因表达式程序设计; 模拟退火; 多种群策略; MPI; 并行算法; 符号回归

**中图分类号:** TP181 **文献标识码:** A **文章编号:** 0372 2112 (2005) 11 2017 05

## Parallel Gene Expression Programming Algorithm Based on Simulated Annealing Method

JIANG Siwei, CAI Zhihua, ZENG Dan, LI Qu, CHENG Yuanfang

(School of Computer Science, China University of Geosciences, Wuhan, Hubei 430074, China)

**Abstract:** This paper presents a hybrid GEP parallel algorithm, which combines simulated annealing and genetic mechanism to avoid the problem of premature convergence. Meanwhile, it imports multi population strategy and migration operator, and a coarse grained parallel computing on finite number of CPUs is implemented. Experiments on sequence induction, function finding and sunspot prediction show that the algorithm has better search ability and higher stability than basic GEP.

**Key words:** gene expression programming; simulate annealing; multi population strategy; MPI; parallel algorithm; symbolic regression

### 1 引言

回归分析方法及其应用的研究由来已久, 近年来演化计算领域的研究者采用树形结构编码的遗传程序设计 (Genetic Programming) 进行符号回归的研究<sup>[1]</sup>, 并在此基础上根据问题的特点设计出了许多算法来提高回归效率, 这些方法在很多经典问题上取得了很好的效果. 2000 年 Ferreira 提出了一种新的遗传算法——基因表达式编程 (Gene Expression Programming, 简记为 GEP)<sup>[2]</sup>, 并将其应用于回归分析的研究<sup>[2-5]</sup>.

基因表达式编程作为一种基于基因型组 (Genotype) 和表现型组 (Phenotype) 的新的遗传算法, 它结合了 GA 和 GP 的优点. 从编码方式来分析, 遗传算法可以分成三个部分<sup>[4,5]</sup>: (1) 基本的遗传算法 (Genetic Algorithms, 简记为 GA), 其个体编码是定长或变长线性字符串; (2) 遗传规划 (Genetic Programming, 简记为 GP), 其个体编码为长度和形状不同的非线性实体, 即表达式树; (3) 基因表达式编程 (GEP), 其个体为成固定长度的线性串 (基因型组或线性串), 然后被表达成不同长度

和形状的非线性实体 (表现型组或表达式树).

从我们早期在回归分析和预测方面的研究结果<sup>[6-8]</sup>可以看出, 虽然基本的 GEP 可以得到较好的结果, 但是仍然存在早熟收敛的问题, 本文在基本 GEP 的基础上, 将模拟退火<sup>[9,10]</sup> (Simulated Annealing, SA) 的思想引入 GEP, 以提高算法跳出局部最优的能力; 同时采用多种群协同进化策略和总最优个体的迁移操作<sup>[11,12]</sup>, 在 MPI 环境下, 实现了并行 GEP 算法, 以此来提高算法的均衡搜索能力, 进一步优化了 GEP 的性能.

### 2 相关工作

在基因表达式编程中, 染色体通常由等长的多个基因构成, 多基因之间采用连接符 (如“+”) 构成更复杂的表达式树结构<sup>[2,3]</sup>. 基因表达式编程是一种基因型组/表现型组的系统, 其搜索空间为线性串, 其解空间为表达式树, 这种搜索空间和解空间分离的特点保证了 GEP 的性能优于基本的遗传

算法<sup>[2,13,14]</sup>。在基因表达式编程中设计了很多遗传操作,测试表明在这些遗传操作中变异是最有效的遗传算子<sup>[14]</sup>。

为提高遗传算法的性能,很多学者引入了模拟退火的思想,如王雪梅<sup>[9]</sup>提出用改进的按适应度比例选择再生方式作为新生子代的生存策略,王凌<sup>[10]</sup>利用遗传算法和模拟退火算法在优化结构和优化行为上的互补性,提出了具有并行搜索和概率突跳特性的全局高效 GASA 混合策略,吴浩扬<sup>[11]</sup>、管宇<sup>[12]</sup>将多种群策略和迁移算子的思想引入遗传算法中,利用并行计算提高了遗传算法的性能。

### 3 基于模拟退火的并行基因表达式编程算法

#### 3.1 基于模拟退火的 GEP 算法

文献<sup>[2]</sup>提出了基本 GEP 算法的流程:(1) 设定控制参数,选择函数集合,初始化种群;(2) 解析 GEP 基因表达式,评价种群;(3) 利用选择、变异、变换(包括 IS transposition, RIS transposition, Gene transposition)、重组(包括 One point recombination, Two point recombination, Gene recombination)、随机常量变异和随机常量变换生成新种群;(4) 实施最优保存策略;(5) 若达到设定最大代数或计算精度,则进化结束,否则返回步骤(2)。

为保证 GEP 算法的性能,GEP 设计了 9 种遗传操作,但是过多的算子对于遗传算法的性能有很大的影响,对于相同的问题不同的参数设置可能得到截然不同的结果。为降低算法对参数的依赖,一个很直接的办法就是减少算子,在算法设计中我们保留了基本 GEP 中最有效的变异算子<sup>[14]</sup>,仅采用了变异、交叉和附加域变异 3 种遗传操作,将模拟退火的思想与 GEP 算法结合,提出了基于模拟退火的 GEP 算法<sup>[15]</sup>(简记为 GEPSA)。

GEPSA 求全局最大值算法步骤如下:

(1) 初始化控制参数:群体规模  $n$ ,退火初始温度  $T_0$ ,温度冷却参数  $\alpha$ ;

(2) 计算种群个体  $P_i(i=0,1,\dots,n)$  的适应度;

(3) 对现有种群实施如下操作,迭代直至产生出下一代新的群体:

①对父代个体  $Old$ ,以变异概率随机选取某个体对其进行变异操作产生子代个体  $New$ ,并计算它们的适应函数值  $f(Old)$  和  $f(New)$ ,令  $\Delta f = f(Old) - f(New)$ ,若  $\min\{1, \exp(-\Delta f/T_i)\} > random$ ,则接收  $New$  为新子代个体。 $random$  为  $[0, 1]$  的随机数;

②对父代个体  $Old$ ,以交叉操作和附加域变异操作对其进行遗传操作产生子代个体  $New$ ,依 ①中方式接受新子代个体;

(4) 实施降温操作,  $T_{i+1} = T_i * \alpha$ ;

(5) 最优个体保存策略;

(6) 若满足停机准则,算法结束;否则,转(3)。

#### 3.2 基于 MPI 的 GEPSA 并行算法

在 GEPSA 算法的基础上,本文采用多种群协同进化策略和总最优个体的迁移操作优化 GEP 算法的性能,提出了基于 MPI 的 GEPSA 并行算法。关于 MPI 环境设置和局域网的并行

计算,主要参考了文献<sup>[16]</sup>。

多种群策略的主要思想是:每个种群采用不同的遗传概率进行不同目标的迭代搜索,遗传概率设定得较小时,对模式的破坏性较弱,可保留大量遗传信息,适合于局部探索,遗传概率设定较大时,不同的模式组合能力较强,可拼接成更多不同类型的积木块,适应度高的个体的产生概率增加,适合于全局探索;不同的种群拥有不同的遗传参数,某些种群偏向于全局最优,某些种群偏向于局部最优,多种群的协同进化提高了算法的均衡搜索能力<sup>[17]</sup>。

总最优个体迁移算子传递不同种群的“优质”遗传信息,它容易改变子种群的组成情况,打破平衡状态,触发接收子种群的进化。迁移算子在多种群并行遗传算法中起着相当重要的作用,最优解通常出现在迁移操作之后<sup>[12]</sup>。

MPI 环境下 GEPSA 并行算法流程,每个处理器节点  $C_m$  ( $m=1,2,\dots,N$ ) 均执行如下算法:

(1) 随机初始化该计算节点种群  $P$ ,不同的种群拥有不同的控制参数,实现不同的搜索目的;

(2) 计算种群中各个体  $P_i(i=1,2,\dots,n)$  的适应度;

(3) 随机选取父代个体,遗传操作产生子代个体,以 Boltzmann 准则接收子代个体,迭代生成下一代种群。

(4) 实施降温操作;

(5) 搜索种群中最优个体和最差个体,并记忆当前最优个体;

(6) 由计算节点  $C_m$  发送当前最优个体到其它各处理器  $C_j(j=1,2,\dots,N,j \neq m)$ ;  $C_m$  接收其它各处理器  $C_j(j=1,2,\dots,N,j \neq m)$  的当前最优个体,比较计算节点最优个体的适应度值,得到总最优个体所在节点  $\max$ ;

(7) 若  $m = \max$ ,发送  $C_m$  的最优个体;否则,接收  $C_{\max}$  的最优个体,并替换掉最差个体;

(8) 若满足进化终止条件,则退出;否则,转(4)。

## 4 实验和结果

#### 4.1 实验数据

本文中的测试集为序列推理、函数发现和太阳黑子预测,其数据均来自文献<sup>[5]</sup>,通过 100 次运行比较基本 GEP、GEPSA 和并行 GEPSA 的性能。

(1) 序列推理,  $N$  的取值根据序列公式得到:

$$N = 5a_n^4 + 4a_n^3 + 3a_n^2 + 2a_n + 1 \quad (1)$$

其中  $a_n$  为非负整数,采用 10 组测试数据( $a_n = 1, 2, \dots, 10$ ),根据给定的序列推理数据以上算法可以找到吻合的方程,在文中通过寻优成功概率比较算法的性能。

(2) 函数发现,已知的曲线形状为“V”的函数表达式为:

$$y = 4.251 * a^2 + \ln(a^2) + 7.243 * e^a \quad (2)$$

其中是  $[0, 1]$  之间的浮点数,随机得到 20 组测试数据,算法不能得到吻合的方程,在 100 次独立运行的基础上,通过平均最优适应度和平均最优 R Square 来比较算法的性能。

(3) 太阳黑子预测问题,用 Wolfer 太阳黑子序列的观测值(见表 1),其隐含的维数是 10,延时为 1,这样可以得到 90 组测试数据,算法得不到吻合的方程,通过比较最优适应度和平

均最优 R Square 来比较算法的性能。

表 1 Wolfer 太阳黑子序列

101	82	66	35	31	7	20	92
154	125	85	68	38	23	10	24
83	132	131	118	90	67	60	47
41	21	16	6	4	7	14	34
45	43	48	42	28	10	8	2
0	1	5	12	14	35	46	41
30	24	16	7	4	2	8	17
36	50	62	67	71	48	28	8
13	57	122	138	103	86	63	37
24	11	15	40	62	98	124	96
66	64	54	39	21	7	4	23
55	94	96	77	59	44	47	30
16	7	37	74				

4.2 遗传操作和适应度函数的选取

本文算法所采用的遗传操作为:

(1) 交叉. 两个染色体的每个基因在一个随机选择的位置交叉, 形成两个染色体的多点交叉。

(2) 变异. 在基因头部, 任何符号都可以变异成函数符号或者终点; 在基因的尾部, 终点只能变异成终点。由于 GEP 编码方式的特点, 由变异产生的新个体在结构上是正确的。

(3) 附加域变异. 基因中的附加域变异成其它的随机值, 附加域元素的取值范围是[- 1, 1]。

本文算法所采用的适应度函数为基于绝对误差的适应度函数(公式 3)和基于相对误差的适应度函数(公式 4), 其中  $M$  是选择范围,  $C_{(i, j)}$  是染色体个体  $i$  对于适应度样本  $j$  (来自  $C_i$  集合中) 的返回值。而  $T_j$  是适应度样本  $j$  的目标值。如果  $|C_{(i, j)} - T_j|$  (即精度) 小于或者等于 0.01, 则取精度为 0。

$$f_i = \sum_{j=1}^{C_i} (M - |C_{(i, j)} - T_j|) \quad (3)$$

$$f_i = \sum_{j=1}^{C_i} (M - |C_{(i, j)} - T_j| / T_j * 100) \quad (4)$$

对于序列推理问题, 适应度函数选择式(4), 选择范围是 20%, 最大的适应度为  $f_{max} = 200$ ; 对于“V”函数问题, 适应度函数选择公式(4), 选择范围是 100%, 最大的适应度为  $f_{max} = 2000$ ; 对于太阳黑子问题, 适应度函数选择公式(3), 选择范围是 1000%, 最大的适应度为  $f_{max} = 90,000$ 。

文中采用 R Square 检验算法建立的模型, 其中  $y_i$  为实际观测值,  $\bar{y}$  为所有  $y_i$  的平均值,  $\hat{y}_i$  为 GEP 建模得到的拟合数值。R square 的取值范围是 [0, 1], 越大的 R square 表示模型的拟合优越性越高<sup>[7]</sup>, 其中

$$R^2 = SS_{Reg} / SS_{Total} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / \sum_{i=1}^n (y_i - \bar{y})^2 \quad (5)$$

4.3 实验结果及其分析

三组实验数据测试的函数符合集、终结符合集和遗传操作等的设置见表 2。其中第 2、6、10 列是使用了随机常量的基本 GEP 算法结果; 第 3、7、11 列是没有使用随机常量的基本

GEP 算法结果; 第 4、8、12 列是 GEPSA 算法结果; 第 5、9、13 列是基于 MPI 的 GEPSA 并行算法结果。为体现算法的稳定性, 在三个数据集上分别独立运行了 100 次, 在表 2 中我们给出平均最优适应度函数和平均最优 R Square, 实验结果表明: 在同等条件测试下, GESA 算法的结果优于基本 GEP 算法; 基于 MPI 的 GEPSA 并行算法优于 GEPSA 算法。

(1) 对于序列推理的问题, 以上 GEP 算法均可以得到完全吻合的方程, 其中文献[5]中 Candida 采用使用了随机变量基本 GEP 算法和不使用随机变量的基本 GEP 算法, 这两种方法涉及到的遗传算子至少有 7 种, 在 100 次独立运行中找到吻合方程的成功概率分别为 16% 和 81%; GEPSA 算法和 GEP-SA 并行算法虽然只采用了 3 种遗传算子, 但是它们的成功概率达到了 89% 和 92%。

(2) 对于函数发现问题, Candida 采用基本的 GEP 算法得到最优方程适应度 fitness = 1990.23, R 平方检验 R Square = 0.9999313; 使用 GEPSA 算法, 在第 87 次运行第 3897 代得到方程如下:

$$y = [0.068 * \cos(0.538) / a - 0.114] + [\exp(\exp(-0.806 * a - \sin(0.216)))] + [\ln|a| * 0.584 / \cos(\sin a)] + [\ln|1 - \cos a|] + [\exp((\cos(-0.066) + 0.5 + a) / 0.792)]$$

该方程的适应度 fitness = 1995.7, R 平方检验 R-Square = 0.999972; 使用基于 MPI 的 GEPSA 并行算法, 第 63 次运行第 2201 代得到方程如下:

$$y = [\sin(\cos(\ln|a|)) * \cos(0.108 + a)] + [\exp(\cos(\sin((a + 0.53) / \sin(0.866)))] + [\exp(a) + 1.354 * a - 0.160156] + [\cos(\ln|\sin(a - 0.014) - 2a|)] + [0.796 + a / 0.162 + a^2 / 0.102]$$

该方程的适应度 fitness = 1996.14, R 平方检验 R Square = 0.99999。

(3) 对于太阳黑子预测问题, 采用基本 GEP 算法得到最优方程适应度 fitness = 89176.61, R 平方检验 R Square = 0.882831; 使用 EPSA 算法, 在第 71 次运行第 4754 代得到方程如下:

$$y = \left[ j + \frac{b * h}{0.808 * b + e * g} \right] + \left[ \frac{a + c}{a * d + i * j * h} * d * j \right] + \left[ -0.262 * \left[ i + \frac{a + b + e}{e - i} \right] \right]$$

该方程的适应度 fitness = 89386, R 平方检验 R Square = 0.924051; 使用基于 MPI 的 GEPSA 并行算法, 第 97 次运行第 2451 代得到方程如下:

$$y = \left[ j + \frac{a * h}{0.694 * a + e * g} \right] + \left[ \frac{a + d}{a * c + i * j * h} * c * j \right] + \left[ -0.272 * \left[ i + \frac{a + 2j}{e - i} \right] \right]$$

该方程的适应度 fitness = 89394.8, R 平方检验 R Square = 0.928966。

为体现算法的寻优能力, 通过三组数据的比较, 从成功率、最大适应度和 R 平方检验分析, 结果表明: GEPSA 得到的最优方程比基本 GEP 算法好; 基于 MPI 的并行算法得到的最

优方程比 GEPSA 得到的方程好.

表 2 在序列推理、函数发现和太阳黑子预测三个问题上的基本设置

	$SI^C$	$SI$	$SI^*$	$SI^{**}$	$V^C$	$V$	$V^*$	$V^{**}$	$SS^C$	$SS$	$SS^*$	$SS^{**}$
运行次数	100	100	100	100	100	100	100	100	100	100	100	100
最大代数	100	100	100	100	5000	5000	5000	2500	5000	5000	5000	2500
种群个数	1	1	1	2	1	1	1	4	1	1	1	4
种群个体数目	100	100	100	50	100	100	100	50	100	100	100	50
样本个数	10	10	10	10	20	20	20	20	90	90	90	90
函数符集合	+ - * /	+ - * /	+ - * /	+ - * /	+ - * / $LEK \sim SC$	+ - * / $LEK \sim SC$	+ - * / $LEK \sim SC$	+ - * / $LEK \sim SC$	+ - * /	+ - * /	+ - * /	+ - * /
终结符集合	$a, ?$	$a$	$a, ?$	$a, ?$	$a, ?$	$a$	$a, ?$	$a, ?$	$a, ?$	$a$	$a, ?$	$a, ?$
随机常量数组长度	10	—	10	10	10	—	10	10	10	—	10	10
随机常量取值范围	0, 1, 2, 3	—	0, 1, 2, 3	0, 1, 2, 3	$[-1, 1]$	—	$[-1, 1]$	$[-1, 1]$	$[-1, 1]$	—	$[-1, 1]$	$[-1, 1]$
头部长度	6	6	6	6	6	6	6	6	8	8	8	8
基因个数	7	7	7	7	5	5	5	5	3	3	3	3
连接符	+	+	+	+	+	+	+	+	+	+	+	+
染色体长度	140	91	140	140	100	65	100	100	78	51	78	78
变异概率	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044
单点交叉概率	0.3	0.3	—	—	0.3	0.3	—	—	0.3	0.3	—	—
两点交叉概率	0.3	0.3	—	—	0.3	0.3	—	—	0.3	0.3	—	—
多点交叉概率	—	—	0.3	0.3	—	—	0.3	0.3	—	—	0.3	0.3
基因重组概率	0.1	0.1	—	—	0.1	0.1	—	—	0.1	0.1	—	—
IS 元素重组概率	0.1	0.1	—	—	0.1	0.1	—	—	0.1	0.1	—	—
IS 元素长度	1, 2, 3	1, 2, 3	—	—	1, 2, 3	1, 2, 3	—	—	1, 2, 3	1, 2, 3	—	—
RIS 元素重组概率	0.1	0.1	—	—	0.1	0.1	—	—	0.1	0.1	—	—
RIS 元素长度	1, 2, 3	1, 2, 3	—	—	1, 2, 3	1, 2, 3	—	—	1, 2, 3	1, 2, 3	—	—
基因重组概率	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
随机常量变异概率	0.01	—	0.01	0.01	0.01	—	0.01	0.01	0.01	—	0.01	0.01
Dc 域重组概率	0.1	—	—	—	0.1	—	—	—	0.1	—	—	—
Dc 域长度	1, 2, 3	—	—	—	1, 2, 3	—	—	—	1, 2, 3	—	—	—
初始温度	—	—	20	20	—	—	50	50	—	—	50	50
温度变化常数	—	—	0.95	0.95	—	—	0.8	0.8	—	—	0.8	0.8
选择范围	20%	20%	20%	20%	100%	100%	100%	100%	1000%	1000%	1000%	1000%
精度	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
平均最优适应度	179.83	197.23	197.72	199.69	1914.8	1931.8	1972.5	1975.3	86215	89033	89166	89181
平均最优 R 平方	0.9776	0.9994	0.9999	0.9999	0.9573	0.9953	0.9966	0.9974	0.7134	0.8119	0.8751	0.8839
成功率	16%	81%	89%	92%	—	—	—	—	—	—	—	—

注: 带“C”表示使用了随机常量的基本 GEP 算法、“\*”表示使用了 GEPSA 算法、“\*\*”表示使用了基于 MPI 的 GEPSA 算法的结果.

## 5 结论

GEP 的特殊编码既克服了 GAs 和 GP 的缺点, 又保留了 GA 和 GP 的优点, 它将线性编码与树型结构的有机结合, 为遗传算法开辟了一种新的思路.

本文将模拟退火和多种群策略引入 GEP 算法中, 提出了一种基于模拟退火的并行基因表达式编程算法, 通过测试序列推理、函数发现和太阳黑子数据三个实例, 100 次独立运行的结果表明: 不仅在最优方程的检验中, GEPSA 并行算法比基本 GEP 算法更好; 而且在平均最优适应度和平均最优 R 平方检验上, GEPSA 并行算法优于基本 GEP 算法. 因此改进的算

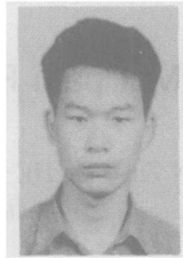
法比基本 GEP 算法具有更强的寻优能力和更高的稳定性.

参考文献:

- [1] Koza J R. Genetic Programming: On the Programming of Computers by Means of Natural Selection [M]. Cambridge, MA: MIT Press, 1992.
- [2] Ferreira C. Gene expression programming: A new adaptive algorithm for solving problems [J]. Complex Systems, 2001, 13 (2): 87-129.
- [3] Ferreira C. 2001. Gene expression programming in problem solving [A]. invited tutorial of the 6th Online World Confer-

- ence on Soft Computing in Industrial Applications [ C ] . Berlin: 2001, 9: 10- 24.
- [ 4 ] Ferreira C. 2002. Genetic representation and genetic neutrality in gene expression programming[ J ] . Advances in Complex Systems, 2002, 5(4) : 389- 408.
- [ 5 ] Ferreira C. Function finding and the creation of numerical constants in gene expression programming[ A ] . the 7th Online World Conference on Soft Computing in Industrial Applications[ C ] . England: September 23 October 4, 2002.
- [ 6 ] 李曲, 蔡之华, 朱莉, 赵云胜. 基因表达式程序设计方法在采煤工作面瓦斯涌出量预测中的应用[ J ] . 应用基础与工程科学学报. 2004, 2(1) : 49- 54.
- [ 7 ] 李曲, 蔡之华, 蒋思伟, 朱莉, 基因表达式程序在预测中的应用研究[ A ] . , 第五届全球智能控制与自动化大会[ C ] . 中国杭州: 2004. 6.
- [ 8 ] Cai Zhihua, Li Qu, Jiang Siwei. Symbolic regression based on GEP and its application in predicting the amount of gas emitted from coal face[ A ] . Proceedings of the 2004 international symposium on safety science and technology [ C ] . China, Shanghai: Vol. IV. Science Press, 2004. 637- 641.
- [ 9 ] 王雪梅, 王义和. 模拟退火算法与遗传算法的结合[ J ] . 计算机学报, 1997, 20(4) : 381- 384.
- [ 10 ] 王凌, 郑大钟. 邻域搜索算法的统一结构和混合优化策略[ J ] . 清华大学学报( 自然科学版) , 2000, 40(9) : 125- 128.
- [ 11 ] 吴浩扬, 常炳国, 朱长纯, 刘君华. 基于模拟退火机制的多种群并行遗传算法[ J ] . 软件学报, 2000, 11(3) : 416- 420.
- [ 12 ] 管宇, 徐宝文. 基于模式迁移策略的并行遗传算法[ J ] . 计算机学报, 2003, 26(3) : 294- 301.
- [ 13 ] Banzhaf W. Genotype phenotype mapping and Neutral Variation A Case Study in Genetic Programming[ A ] . In Y Davirdor, H- P Schwefel, R MÄnner, eds. , Parallel Problem Solving from Nature III, Vol. 866 of Lecture Notes in Computer Science[ C ] . Jerusalem, Israel: Springer Verlag, 1994.
- [ 14 ] Ferreira C. Mutation, transposition, and recombination: An analysis of the evolutionary dynamics[ A ] . Proceedings of the 6th Joint Conference on Information Sciences, 4th International Workshop on Frontiers in Evolutionary Algorithms [ C ] . USA: 2002. 614- 617.
- [ 15 ] Jiang Siwei, Cai Zhihua, Zeng Dan, Liu Yadong, Li Qu. Gene expression programming based on simulated annealing [ A ] . International Conference on Wireless Communications, Networking and Mobile Computing [ C ] . China, Wuhan: September, 2005. 23- 26.
- [ 16 ] 都志辉. 高性能计算之并行编程技术 MPI 并行程序设计[ M ] . 北京: 清华大学出版社, 2001.

#### 作者简介:



蒋思伟 男, 1980 年 11 月生, 湖北天门人, 中国地质大学计算机学院硕士研究生, 主要研究领域为数据挖掘、基因表达式编程、并行算法和贝叶斯网络。



蔡之华 男, 1964 年 5 月生, 湖北浠水人, 教授, 博士生导师, 主要研究领域为数据挖掘、演化算法和贝叶斯网络. E-mail: zhcai@ uig. edu. cn.