均匀设计法在 GA 欺骗问题中的应用研究

张克　刘永才　关世义　涂震飚

中国航天科工集团第三研究院　　北京

摘 要： 针对遗传算法中的欺骗问题，通过构造最小欺骗问题，利用均匀设计成功地克服了欺骗问题，该方法可以推广到解决遗传算法的其它欺骗问题中。

关键词： 均匀设计；遗传算法；欺骗问题

1 引言

Frantz 认为，常使遗传算法(Genetic Algorithms, GA)从全局最优解发散出去的问题，成为 GA——欺骗(GA-deceptive)问题。所谓的欺骗问题，即构造一个问题，给定一些带欺骗性的初始条件，“迷惑”GA，使其偏离全局最优解。为此，要最大限度地违背积木块假设(由模式理论，一个问题是否能用 GA 求解，取决于问题的编码是否满足积木块假设，满足者用 GA 求解效率高，不满足者效率低、甚至找不到满意解)，即使低阶、短距、高平均适应度的模式生成局部最优点，或者导致算法发散，找不到最优解[3~5]。欺骗问题实际就是要预测给定问题用 GA 求解的难易程度。由文献[3]可知，目前 GA 的欺骗问题研究主要集中在三个方面：设计欺骗函数；理解欺骗函数对 GA 的影响；修改 GA 以解决欺骗函数的影响。

均匀设计是统计试验设计的方法之一，它的数学原理是数论中的一致分布理论，此方法借鉴了“近似分析中的数论方法”这一领域的研究成果，将数论和多元统计相结合，是属于伪蒙特卡罗方法的范畴。它着重在实验范围内考虑试验点均匀散布以求通过最少的实验来获得最多的信息，因而其试验次数比正交设计明显减少，使均匀设计特别适合于多因素多水平的试验和系统模型完全未知的情况。

对于遗传算法的其它欺骗问题，由于参数众多，推导过于复杂利用，因此为了说明问题方便，本文构造了最小欺骗问题，个体中只含 2 个基因，利用均匀设计成功地克服了最小欺骗问题。该方法可以进行推广，特别适合于 n 个基因的情况，这在 GA 欺骗问题研究中是一种崭新的尝试。

2 最小欺骗问题

以下描述最小欺骗问题(minimal deceptive problem)。这里所谓“最小”是指为了造成骗局所需设置的最小问题的规模(即阶数)。

假设有一个由 4 个阶数为 2、有 2 个确定位置的模式构成的集合，各模式具有如下的适应度：

$$* \ * \ * \ 0 \ * \ * \ * \ * \ * \ 0 \ * \qquad\qquad f(00)$$

$$* \; * \; * \; 0 \; * \; * \; * \; * \; * \; 1 \; * \qquad f(01)$$

$$* \; * \; * \; 1 \; * \; * \; * \; * \; * \; 0 \; * \qquad f(10)$$

$$* \; * \; * \; 1 \; * \; * \; * \; * \; * \; 1 \; * \qquad f(11)$$

式中 $f(00)$，$f(01)$，$f(10)$，$f(11)$ 是各模式的平均适应度，并假设为常值。不妨假定 $f(11)$是全局最优值，即有

$$f(11)>f(00)，f(11)>f(01)，f(11)>f(10) \tag{1}$$

现在，设法引入迷惑 GA 的条件。考虑 4 个一阶模式的适应度，即 $f(*\,0)$、$f(*\,1)$、$f(0\,*)$和 $f(1\,*)$。一阶模式的适应度等于其包含的所有二阶模式的适应度的平均值，即有

$$f(*\,0)=[f(00)+f(10)]/2 \tag{2}$$

$$f(*\,1)=[f(01)+f(11)]/2 \tag{3}$$

$$f(0\,*)=[f(00)+f(01)]/2 \tag{4}$$

$$f(1\,*)=[f(10)+f(11)]/2 \tag{5}$$

令包含全局最优解 $f(11)$的一阶模式的适应度小于不包含最优解的一阶模式的适应度，从数学上讲，就是

$$f(0\,*)>f(1\,*)\;，f(*\,0)>f(*\,1) \tag{6}$$

由式(2)～(6)有

$$[f(00)+f(01)]/2>[f(10)+f(11)]/2 \tag{7}$$

$$[f(00)+f(10)]/2>[f(01)+f(11)]/2 \tag{8}$$

式(7)、(8)给出所谓的“骗”条件。同时可以看出，式(7)和(8)并不能同时成立，否则就有 $f(00)>f(11)$，从而违背了 $f(11)$是全局最优解的假定。不失一般性，不妨假定式(7)成立。由此，通过一个全局条件($f(11)$为全局最优值)和一个“骗”条件($f(0\,*)>f(1\,*)$)，就确定一个欺骗问题。

将上述各适应值按 $f(00)$进行规一化处理如下

$$r=f(11)/f(00)，c=f(01)/f(00)，c'=f(10)/f(00) \tag{9}$$

则全局条件式(1)可以表示为

$$r>1，r>c，r>c' \tag{10}$$

于是，“骗”条件(7)可以表示为

$$r<1+c-c' \tag{11}$$

由式(10)(11)，可以推出

$$c'<1，c'<c \tag{12}$$

由式(6)(7)(12)可以看出，存在如下两类欺骗问题：

类型 1：$f(11) > f(01) > f(00) > f(10)$　$(c>1)$

类型 2：$f(11) > f(00) > f(01)$　　　$(c\leq 1)$

两类问题中的适应度是两个自变量(确定位 1、2)的函数。通过研究发现，一阶问题(仅有一个确定位) 中不可能存在欺骗问题，因为无法找到与全局条件相适应的骗条件。因此，上述 2 阶欺骗问题是可能存在的最小欺骗问题。

3 均匀设计在解决欺骗问题中的应用

对于最小欺骗问题，即种群的个体值包含两个因素。相应地，用均匀设计(UD)时只考虑两个因素 A、B 和两个因素的交互项 A×B。由于每个因素的水平数只有 0 或 1，因素数大于水平数，无法很好利用均匀设计，也体现不出均匀设计的优势。为了解决这一矛盾，借鉴了正交设计的思想，由于正交设计可视为均匀设计的特例[1,8]。因此，经过分析可以构造出的因素水平表见表 1，均匀设计表见表 2。

表 1　因素水平表

因素 水平	A	B	A×B
1	0	0	0
2	1	1	1
3	0	0	0
4	1	1	1

表 2　$U_4(4^3)$的均匀设计表

	1	2	3
1	1	2	2
2	2	1	4
3	3	4	3
4	4	3	1

选择好水平表和均匀设计表后，就可以安排试验。具体见表 3。

表 3 求解 GA 欺骗问题的均匀设计试验方案

因素及列号 试验号	A	B	A×B	试验结果
1	1(0)	2(1)	2(1)	$f(01)$
2	2(1)	1(0)	4(1)	$f(10)$

3	3(0)	3(0)	3(0)	$f(00)$
4	4(1)	4(1)	1(0)	$f(11)$
K_0	$(f(01)+f(00))/2$	$(f(10)+f(00))/2$	$(f(00)+f(11))/2$	
K_1	$(f(10)+f(11))/2$	$(f(01)+f(11))/2$	$(f(01)+f(10))/2$	
极 差 R	R_A	R_B	R_{AB}	

其中 K_0 行对应于因素取 0 水平的有关试验的效果平均值，K_1 行对应于因素取 1 水平的有关试验的效果平均值，极差 R 是 K_0 与 K_1 差的绝对值。

$$R_A = |f(01) + f(00) - f(10) - f(11)|/2$$

$$R_B = |f(10) + f(00) - f(01) - f(11)|/2$$

$$R_{AB} = |f(00) + f(11) - f(01) - f(10)|/2$$

极差越大表明对应的因素或交互作用项对实验效果的影响就越强。

(1) 类型 1 的欺骗问题

$$R_{AB} = |f(00) + f(11) - f(01) - f(10)|/2 = |(f(00) - f(10)) + (f(11) - f(01))|/2$$

$$R_A = |f(01) + f(00) - f(10) - f(11)|/2 = |(f(11) - f(01)) - (f(00) - f(10))|/2$$

因为 $f(11) - f(01) > 0$，$f(00) - f(10) > 0$，坪盼 $R_{AB} > R_A$。

$$R_{AB} = |f(00) + f(11) - f(01) - f(10)|/2 = |(f(11) - f(10)) - (f(01) - f(00))|/2$$

$$R_B = |f(10) + f(00) - f(01) - f(11)|/2 = |(f(11) - f(10)) + (f(01) - f(00))|/2$$

因为 $f(11) - f(10) > 0$，$f(01) - f(00) > 0$，坪盼 $R_{AB} < R_B$

可见，因素 B、交互作用项 A×B 起主要作用。对因素 B 来说，$K_1 = (f(01) + f(11))/2 > (f(01) + f(00))/2 = K_0$，所以因素 B 取 1 水平为好(即对应于 1 、 4 号试验)；对交互作用项 A×B 来说，$K_0 = (f(00) + f(11))/2 > (f(01) + f(10))/2 = K_1$，所以交互作用项 A×B 取 0 水平为好(即对应于 3、4 号试验)。综合判断后可得结论：4 号试验(即 1 1 模式)为最优方案，故对类型 1 欺骗问题而言，欺骗被克服了。

(2) 类型 2 的欺骗问题

$$R_{AB} = |f(00) + f(11) - f(01) - f(10)|/2 = |(f(00) - f(01)) + (f(11) - f(10))|/2$$

$$R_B = \left| f(10) + f(00) - f(01) - f(11) \right| / 2 = \left| (f(00) - f(01)) - (f(11) - f(10)) \right| / 2$$

因为 $f(11) - f(10) > 0$，$f(00) - f(01) > 0$， 坪眕$R_{AB} > R_B$

$$R_A = \left| f(01) + f(00) - f(10) - f(11) \right| / 2 = \left| (f(11) - f(00)) - (f(01) - f(10)) \right| / 2$$

$$R_B = \left| f(10) + f(00) - f(01) - f(11) \right| / 2 = \left| (f(11) - f(00)) + (f(01) - f(10)) \right| / 2$$

因为 $f(11) - f(00) > 0$，$f(01) - f(10) > 0$， 坪眕$R_B > R_A$

可见，对于类型 2 问题，仍然是交互作用项 A×B 和因素 B 起主要作用。类似前面的推理可知，交互作用项 A×B 应取 0 水平(即对应于 3、4 号试验)，因素 B 应取 1 水平(即对应于 1、4 号试验)。综合判断可得结论，4 号试验(即 11 模式)为最优方案，故对类型 2 的欺骗问题而言，欺骗也被克服了。

4 结论

以上就个体只包含两个基因的情况进行讨论。一般地，当个体包含 n 个基因时，对应的均匀试验设计法就要包含 n 个因素、n(n-1)/2 个两两因素交互作用项。类似于上述步骤，可以对 n 个因素中的任意两个因素及其对应的一个两两因素交互作用项进行讨论，同理可以证明均匀设计法能解决两种类型的欺骗问题，给出正确结果。

综上所述可见，均匀试验设计法可以用来解决一般遗传算法中的欺骗问题。

参考文献

1 方开泰，马长兴．正交与均匀实验设计．科学出版社. 2001

2 刘永才，王建民．均匀设计在飞航导弹研制中的应用．战术导弹技术. 2001, (4):1~4

3 陈国良，王煦法等．遗传算法及其应用.人民邮电出版社, 1996:5～89

4 Feng Q X, Francesco P. The Oretical Analysis of Evolutionary Algorithms with an Infinite Population Sizein Continuous Space Part and : Basic Properties of Selection and Mutation. IEEE Trans Neural Networks, 1994, 5(1):102~129

5 Rudolph G. Convergence Analysis of Canonical Genetic Algorithms. IEEE Trans Neural Networks, 1994, 5(1): 96~101

6 陈魁.试验设计与分析．清华大学出版社，1996:90~123

7 张克，刘永才,关世义，强文义．均匀设计在遗传算法中的应用研究．战术导弹技术. 2003(5): 35~40

8 张克．面向巡航导弹任务规划的攻防对抗研究．博士学位论文. 2002：103~115