

基于遗传算法的工程项目评标

段爱玲¹, 王伟², 程凤娟¹

(1. 河南工业大学信息科学与工程学院, 郑州 450001; 2. 郑州牧业工程高等专科学校信息工程系, 郑州 450058)

摘要: 评标是工程项目投标中的重要环节, 利用遗传算法进行评标, 其难点在于缺乏合理可行的适应度函数设计方法。在给出遗传算法编码规则的基础上, 引入层次分析法(AHP)进行适应度函数设计, 通过实例计算, 阐述利用 AHP 方法进行适应度函数设计的全过程。结果证明, 该算法可有效应用于工程项目评标。

关键词: 遗传算法; 适应度函数; 工程项目; 评标

Engineering Project Bidding Evaluation Based on Genetic Algorithm

DUAN Ai-ling¹, WANG Wei², CHENG Feng-juan¹

(1. College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001;

2. Department of Information Engineering, Zhengzhou College of Animal Husbandry Engineering, Zhengzhou 450058)

【Abstract】Bidding evaluation is an important part during engineering projects. The difficulty using genetic algorithm is lack of the fitness function design methodology that is proper and feasible. On the basis of coding rule of genetic algorithm, this paper leads the Analytic Hierarchy Process (AHP) to the design of fitness function. It explains all the process of making the design of fitness function in the use of AHP method through the example calculation. Results show that the algorithm can be effectively applied to the bidding evaluation of engineering projects.

【Key words】 genetic algorithm; fitness function; engineering project; bidding evaluation

评标是投标制中的重要环节, 如何选择合适的评标方法, 关系到工程项目的最大价值能否得以实现。科学、有效、公正、合理的评标方法在工程项目招标中的成功运用, 对于保证设备质量将起到重要作用。因此, 本文提出一种全新的基于遗传算法的求解方案, 能够明显提高求解问题的速度和精度。

1 问题的提出

实际的工程项目种类繁多, 本文针对设备采购的评标问题, 采用遗传算法进行评标。遗传算法在应用中最关键的问题是串的编码方式、适应度函数和自身参数的设计^[1]。遗传算法把问题的解表示成“染色体”, 在算法中即二进制编码的串。并且, 在执行遗传算法之前给出一群染色体, 即假设解。将这些假设解置于问题的环境中, 并按适者生存的原则, 从中选择出较适应环境的染色体进行复制, 再通过交叉、变异过程产生更适应环境的新一代染色体群。这样, 一代一代地进化, 最后就会收敛到最适应环境的一个染色体上, 它就是问题的最优解^[2-3]。层次分析法(Analytic Hierarchy Process, AHP)是把复杂问题中的各因素划分成相关联的有序层次, 使之条理化, 根据一定客观的判断, 对每一层次中每 2 个元素相对重要性给出定量表示, 从而确定出全部元素的权重。该方法适合于标书指标结构复杂且缺乏必要数据的评标条件。

2 染色体编码

一般把问题的各种参数用二进制编码, 构成子串, 然后把子串拼接构成染色体串, 串长度及编码形式对算法收敛影响极大, 编码原则的完备性应全面地考虑厂商提供的产品信息, 以及厂商情况。为了操作的方便, 利用一维染色体中的二进制编码方法来对厂商提供的信息进行编码, 确定客户提

供信息, 计算机评标的厂商参数如表 1 所示。

表 1 计算机评标的厂商参数

	CRTcost	CPUCost	MMcost	SCcost	CPU	MM	SC	PREST	DEM
CLIENT1	856	530	456	540	2.5	256	64	好	1
CLIENT2	906	610	550	583	2.6	256	128	良	1
CLIENT3	966	588	546	526	2.8	512	64	好	0

对 3 个投标商的投标信息进行编码如下:

CLIENT1: 01 00 00 00 01 00 01 11 1

CLIENT2: 10 10 01 01 01 00 10 01 1

CLIENT3: 11 01 01 00 11 01 01 11 0

3 适应度函数设计

适应度函数的设计是遗传算法中一个至关重要的技术环节, 只有当遗传函数能科学的计算出每个染色体的适应度值时, 才能保证遗传操作准确。因此, 适应度函数直接影响到遗传算法的收敛速度以及能否找到全局最优解, 根据考虑的问题空间信息来构造一个类似 $f(x)$ 的函数, 以此来计算每个染色体的适应度值^[2-3], 但本文中使用了 AHP 算法来求解每个个体的适应度值, 把每个个体的适应度值在一起求解, 在考虑各个信息参数的同时, 也考虑到同其他个体的信息, 以此求得适应度值。

在利用 AHP 分析法进行适应度函数设计中, 这里只要算

基金项目: 河南省自然科学基金资助项目“基于遗传算法的工程项目评标研究”(200648)

作者简介: 段爱玲(1964 -), 女, 副教授, 主研方向: 人工智能; 王伟, 博士; 程凤娟, 硕士

收稿日期: 2009-02-26 **E-mail:** duanailing@126.com

出每代中群中的每个个体的适应度值，在层次单排序和层次总排序中将略去一致性检验，即认为判断矩阵在经协商确定后就不可改变，以省去一些繁琐操作。

3.1 层次结构模型的构建

在对计算机的评标过程中，目标层就是中标单位，措施层就使投标商，对于准则层要进行细分，其中准则层中为价格、性能和厂商情况。对于每一项在自准则层中又进行细分，最终建立的层次结构模型。

3.2 判断矩阵的构造

判断矩阵表示针对上一层次中的某些元素而言，评定该层次中各有关元素的相对重要性。在判断矩阵的构造过程中，需要综合各方的意见，从而制定各个因素在评标过程中所占的比重，如对于价格、性能、厂商情况，就需要学校计算机使用单位，社会专家根据协商以确定三者的相对重要性。对于中标单位中价格、性能和厂商情况三者构造判断矩阵：

A	B ₁	B ₂	B ₃
B ₁	1	2	3
B ₂	1/2	1	2
B ₃	1/2	1/2	1

在用遗传算法的处理过程中，判断矩阵是要改变的，要根据每代的种群中的染色体来构造判断矩阵，所以这里要制定相应的规定来构造判断矩阵。在构造最底层判断矩阵时，只能根据每个 CLIENT 的染色体编码来构造。要将这 2 位基因转化为十进制数，以此来构造判断矩阵，具体操作如下：

CLIENT1 :	01 -----	0 × 2 + 1 × 2 ⁰ + 1 = 2
CLIENT2 :	01 -----	0 × 2 + 1 × 2 ⁰ + 1 = 2
CLIENT3 :	11 -----	1 × 2 + 1 × 2 ⁰ + 1 = 4

在上面的计算中每个式子都进行了加 1 操作，这主要是为了对于出现基因为 00 这种情况的一种处理机制，当然也可以采用其他方法。

例如，构造对于 CPU 性能这项参数各个 CLIENT 的相对比重，判断矩阵如下：

C5	CLIENT1	CLIENT2	CLIENT3
CLIENT1	1	1	1/2
CLIENT2	1	1	1/2
CLIENT3	2	2	1

根据上面介绍的方法同理构造出各个 CLIENT 对于其他参数的判断矩阵。在上面构造判断矩阵的过程中必须根据每个 CLIENT 的染色体编码进行，这样就给操作带来了一些限制，在此采用的是对每个参数的相应位置上的基因转化为十进制数进行，其实在此过程中存在着一定的误差，如对于 CPU 性能有 2 个 CLIENT 提供的信息分别为 1.3 GHz, 1.8 GHz, 经过编码后都为 00，这样其在构造的判断矩阵中所对应的比值就为 1，即这 2 个 CLIENT 的这项参数信息将按相同来处理，但在实际中还是存在一定的差别的，若直接用提供的 2 个参数来构造判断矩阵则精确度就会有很大提高，但这在遗传操作中的可行性将受到很大限制。因此，这种操作方式具有一定的误差，但在评标过程中考虑到相关参数越多，越有利于遗传操作。

3.3 次单排序

在层次单排序中将确定本层次和上层中的某些元素有联系的各元素重要性次序的权重值，这就需要根据所构造的判断矩阵求出其最大特征值和对应特征向量。在求最大特征向量的过程中，本文采用方根法。

计算判断矩阵每一行元素的乘积：

$$M_i = \prod_{j=1}^n b_{ij}, i = 1, 2, \dots, n$$

计算 M_i 的 n 次方根：

$$\bar{W} = \sqrt[n]{M_i}, i = 1, 2, \dots, n$$

将向量 $\bar{W} = [\bar{W}_1, \bar{W}_2, \dots, \bar{W}_n]^T$ 归一化，则 $W = [W_1, W_2, \dots, W_n]^T$ 即为所求的特征向量。

计算最大特征根：

$$\lambda_{\max} = \sum_{i=1}^n \frac{(AW)_i}{nW_i}$$

其中， $(AW)_i$ 表示向量 AW 的第 i 个分量。

本文只针对判断矩阵 A 进行求解，其他判断矩阵的求解方式相同。

(1) 求判断矩阵每一行元素的乘积，放入数组 $M[i]$ 中，得：

$$M[0] = 1 \times 2 \times 3 = 6$$

$$M[1] = 1/2 \times 1 \times 2 = 1$$

$$M[2] = 1/3 \times 1/2 \times 1 = 1/6$$

(2) 计算 $M[i]$ 的 3 次方根结果放入 $H[i]$ 中，得：

$$H[0] = \sqrt[3]{M[0]} = 1.817$$

$$H[1] = \sqrt[3]{M[1]} = 1.000$$

$$H[2] = \sqrt[3]{M[2]} = 0.550$$

(3) 将向量 $W = \{H[0], H[1], H[2]\}$ 归一化，得：

$$M = H[0] + H[1] + H[2] = 3.367$$

$$W[0] = H[0]/M = 0.540$$

$$W[1] = H[1]/M = 0.297$$

$$W[2] = H[2]/M = 0.163$$

则数组 W 中的每一项即对应这判断矩阵 A 中各项参数的比重，即对应与影响评标结果的 3 项指标价格、性能和厂商情况的比重值。可以清楚地看出，价格占评标的比重为 54%，性能占的比重为 29.7%，而厂商情况占的比重为 16.3%。

根据以上计算方法，就可以计算出其他各层的单排序结果。这个操作是用一个算法解决的，只须将判断矩阵作为参数传入，就可以计算出相应的最大特征值和对应的特征向量。

3.4 次总排序

在层次总排序中，就要根据每一层所有的层次单排序结果计算出针对上一层，本层中所有元素的重要性权值，需要从上到下，逐层进行，直至措施层，就可以计算出每个 CLIENT 的权重值，这样就求出了每个染色体的适应度值。

经详细计算得出的层次单排序结果如下：

(1) 在价格层次中，各参数所占的比重分别为：CRTcost 占 12.7%，CPUcost 占 50.8%，MMcost 占 21.4%，Scost 占 15.1%。

(2) 在性能层次中，各参数所占的比重分别为：CPU 占 66.1%，MM 占 20.8%，SC 占 13.1%。

(3) 在厂商情况层次中，各参数所占的比重分别为：信誉度占 49.3%，国内占 31.1%，国外占 19.6%。

上面的这些层次单排序结果，只要判断矩阵确定后，再利用遗传算法进行操作的过程中，就不会改变了。

表 2 给出针对每项具体的参数给出例子中 3 个 CLIENT 所占的比重，在此之前应先求出针对每项参数 3 个 CLIENT 的判断矩阵。求出所有层次单排序结果后，就可以计算层次总排序了。求子准则层中各因素关于总目标的权重，即求子准则层各因素的层次总排序权重，如表 3 所示。

表 2 每项参数所占的比重

参数	所占的比重/(%)		
	CLIENT1	CLIENT2	CLIENT3
CRTcost	46.2	30.8	23.0
CPUcost	65.1	8.0	26.9
MMcost	50.0	25.0	25.0
SCcost	40.0	20.0	40.0
CPU	25.0	25.0	50.0
MM	25.0	25.0	50.0
SC	28.6	42.8	28.6
信誉度	40.0	20.0	40.0
国内	40.0	40.0	20.0
国外	25.0	25.0	50.0

表 3 子准则层各因素的层次总排序权重

参数	层次总排序权重
CRTcost	0.068 580
CPUcost	0.274 320
MMcost	0.115 560
SCcost	0.081 540
CPU	0.196 317
MM	0.061 776
SC	0.038 907
信誉度	0.080 359
国内	0.050 693
国外	0.031 948

由此可以求出所有厂商对于总目标的权重值,即该个体的适应度值。CLIENT1 的适应度值为 0.428 733 732, CLIENT2 的适应度值为 0.205 790 686, CLIENT3 的适应度值为 0.349 501 582。可以看出,在这 3 个投标商中,CLIENT1 为最优,其次为 CLIENT3,最后为 CLIENT2。

4 结束语

本文主要针对遗传算法中的编码问题和适应度函数的设计展开讨论,基本解决了评标问题中的编码问题和适应度函数的理论问题,并在其中合理地利用了 AHP 加权算法解决了适应度函数的设计问题^[4-5]。本文给出了具体的例子,通过详细的数学计算验证了算法的可行性,但在编码和适应度函数的设计方面还不完善,这是下一步的研究方向。

参考文献

- [1] Su Limin, Hou Chaozhen, Dai Zhongjian, et al. The Application of Improved Genetic Algorithm in Network Fault Diagnosis Expert System[J]. Journal of Beijing Institute of Technology, 2003, 12(3): 225-229.
- [2] Tanaka F, Nakatsugawa M, Yamamoto M, et al. Developing Support System for Sequence Design in DNA Computing[C]//Proc. of the 7th International Workshop on DNA-based Computers. Tampa, FL, USA: [s. n.], 2001.
- [3] 王东龙, 李茂青. 基于遗传算法的数据挖掘技术应用[J]. 南昌大学学报, 2005, 27(1): 81-84.
- [4] 刘文斌. DNA 计算中的编码问题及模型研究[D]. 武汉: 华中科技大学, 2003.
- [5] 杨青. 基于遗传算法的试题库自动组卷问题的研究[J]. 济南大学学报: 自然科学版, 2004, 18(3): 228-231.

编辑 顾姣健

(上接第 188 页)

表 1 不同融合方法下多类 SVM 的识别率 (%)

融合方法	PCA	KPCA	LDA	2DPCA
OAASVM	94.3	91.4	87.9	92.1
OAOSVM	93.6	90.0	87.1	89.3
MAX	95.0	90.0	87.9	92.1
MIN	94.3	89.3	88.6	93.6
MEDIAN	94.3	92.1	87.1	93.6
MEAN	94.3	92.1	87.9	94.3
PROD	95.0	90.0	88.6	93.6
VOTE	93.6	92.9	89.3	93.6
DTS1	95.0	92.9	88.6	93.6
DTS2	94.3	92.1	90.0	94.3
DTI1	95.0	93.6	89.3	94.3
DTI2	95.7	94.3	90.0	92.9
DTE	95.0	91.4	89.3	93.6

表 1 列出了各种融合途径实现的多类 SVM 在 JAFFE 库上实现的人脸表情识别结果,其中包括了 2 种典型的多类 SVM,即 OAASVM 和 OAOSVM。采取 4 种特征提取方法进行对比实验,即主元分析(PCA)、核主元分析(KPCA)、线性鉴别分析(LDA)和二维主元分析(2DPCA)。核函数选用高斯径向基函数,取 $\sigma=3$,支持向量机惩罚因子取值 $C=100$ 。JAFFE 库共有 213 幅图像,取每人每种表情中的一幅,共 $10 \times 7 = 70$ 幅图像作训练集,其他 143 幅图像构成测试集。

从表 1 可以看出,大部分融合途径实现的多类 SVM 获得了比 OAASVM 和 OAOSVM 更好的性能,而一些融合方法

的性能则低于 OAASVM 和 OAOSVM,但可以确定,通过融合方法能实现多类 SVM 功能。

5 结束语

SVM 是模式识别中性能优越的分类器算法。本文提出一种基于分类器融合算法的多类 SVM 构成方法。通过融合途径实现多类 SVM 的思路可以推广至任何需要从两类问题扩展为多类问题的模式识别方法。

参考文献

- [1] 李华庆. 支持向量机及其在人脸识别中的应用研究[D]. 上海: 上海交通大学, 2006.
- [2] Kuncheva L I. Decision Templates for Multiple Classifier Fusion: An Experimental Comparison[J]. Pattern Recognition, 2001, 34(2): 299-314.
- [3] 王珏, 周志华, 周傲英. 机器学习及应用[M]. 北京: 清华大学出版社, 2006.
- [4] Alexandru N M, Caruana R. Predicting Good Probabilities with Supervised Learning[C]//Proc. of ICML'05. Bonn, Germany: [s. n.], 2005: 625-632.
- [5] 刘晓旻, 谭华春, 章毓晋. 人脸表情识别研究的新进展[J]. 中国图象图形学报, 2006, 11(10): 1359-1368.

编辑 陈晖