

遗传算法参数设置及其在负荷建模中应用

金 群,李欣然

(湖南大学 电气与信息工程学院,湖南 长沙 410082)

摘要:以基于实测的电力系统综合负荷建模为应用对象,探讨遗传算法的运行机理,分析遗传算子的不同搜索能力。指出决定遗传算法性能的关键因素是种群多样性,得出了种群多样性与算法参数的关联约束。从理论分析上给出遗传参数的设定规则,深入研究遗传算法中种群规模、交叉、变异概率及其控制策略,以及初始种群参数区间等遗传算法关键操作参数对算法性能的影响规律,给出合理的种群规模和参数初始区间,提出与群体进化程度指标相关的自适应调整交叉概率和变异概率策略。研究结果表明,合理的参数组合是挖掘遗传算法潜能的关键,可提高遗传算法运行效率、克服早熟及尽量减小模型参数分散性。

关键词:电力系统; 负荷建模; 遗传算法; 控制参数; 参数区间

中图分类号: TM 714

文献标识码: A

文章编号: 1006-6047(2006)05-0023-05

0 引言

遗传算法(GA)模拟生物体的进化过程,是一种基于自然选择与遗传机理的随机搜索算法。GA 具有将复杂的非线性问题经过有效搜索和动态演化而达到优化状态的特性,在求解复杂优化问题上具有巨大的潜在优势。这种优势主要表现在 2 个方面:一是它对目标函数没有任何解析性质的要求;再者,它具有理论上的全局寻优能力以保证在整个解空间搜索到全局最优解^[1]。

理论上,GA 的这种优良性能能够很好地解决电力负荷建模中准则函数难以解析描述以及模型参数分散性的难题,然而在实践应用中 GA 的这种潜在优势能否得以发挥取决于诸多因素,其中 GA 控制参数的选取是关键。在 GA 的搜索过程中每个参数都表现为多方面的功效,不同的参数组合对 GA 性能的影响非常复杂^[1]。

实践应用表明,对于给定的一类应用对象,如何设置 GA 参数是一个很重要的问题,其规律性把握更是一项十分困难的工作。当前遗传算子的改进和参数的选择都是基于经验获得的,并根据具体问题的仿真结果检验,没有从理论上给出指导性的结论和分析方法。

本文在将 GA 具体应用于基于实测的电力系统综合负荷建模的同时,深入研究遗传算法的运行机理,指出决定 GA 搜索能力的关键因素是种群多样性,并得出其与算法参数的关联性规律,从理论分析上给出了遗传参数的设定规则。

1 操作参数对负荷建模的影响

GA 作为一种具有全局寻优能力的随机性算法,其性能由在搜索空间进行的深度搜索和广度搜索的平衡决定,而这种平衡受种群规模、最大遗传代数、交叉和变异概率诸多参数影响^[1-2]。

GA 以群体和遗传算子为基础,形成一种迭代搜索算法。群体分析和设计的基本内容一般是指群体规模和多样性,在种群规模一定的情况下,种群多样性越大,它就越有可能产生出更优子代,而一旦种群丧失多样性,其个体间越相互趋同,群体缺乏有效等位基因,在遗传算子的作用下不能生成高阶竞争模式,则过早收敛现象就越容易发生^[3]。对于遗传运算过程而言,种群的多样性保证了遗传算子正常发挥进化和重组能力。其中,选择算子的作用减弱群体的多样性,交叉算子保持群体多样性不变,变异算子提高群体的多样性。

因此,种群的多样性决定了 GA 的搜索能力,进化种群始终保持多样性是遗传算法有效运行的前提条件与关键因素。研究遗传算法的参数对负荷建模结果的影响必须与群体多样性联系起来。

1.1 种群规模

种群规模(N)对种群多样性、模式生成、计算量都有显著影响。 N 太小会造成有效等位基因先天缺乏,生成具有竞争力的高阶模式的可能性极小;为了让初始解分布均匀,在计算量允许的情况下, N 应尽量大些以保证种群的多样性。但 N 也不可太大, N 过大将使个体适应性评价的计算量急剧增加,收敛速度显著降低。有关 GA 专著推荐 N 经验值为 20~160^[1]。显然,对于一个具体的工程优化问题,在如此大的 N 取值范围内,GA 性能变化将非常大,且其变化规律无

疑与具体应用对象有关。本文结合电力系统综合负荷建模实践,研究种群规模对 GA 综合性能的影响。

在将遗传算法应用于负荷建模时,综合负荷采用 3 阶感应电动机并联恒定阻抗模型^[4],建模样本取自实际变电站的现场实测负荷特性记录,研究结果如图 1 所示,其中 t 为负荷建模的参数优化辨识时间; g 为迭代至收敛的进化代数; e 为负荷实测响应与模型响应之间的综合误差。

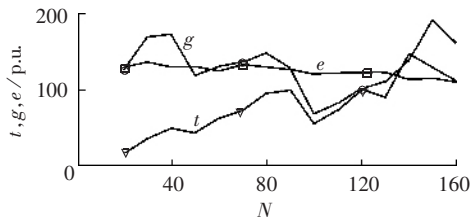


图 1 N 的变化对负荷建模的影响

Fig.1 Effect of N on power load modeling

限于篇幅,建模过程及其模型描述从略^[5]。图中给出了反映 GA 性能的关键参数随种群规模 N 的变化规律。 e 由式(1)定义。式(1)中 L 是实测样本数据长度; $P(k)$, $Q(k)$ 为负荷实测响应; $P_m(k)$, $Q_m(k)$ 为与负荷实测响应相对应的模型响应。由式(1)可知, e 的大小反映了模型响应对实测响应的各对应数据点的平均拟合误差。

$$e = \frac{1}{2L} \sqrt{\sum_{k=1}^L [(P_m(k) - P(k))^2 + (Q_m(k) - Q(k))^2]} \quad (1)$$

由图 1 可知,随着群体规模的增大,收敛代数及迭代时间显著变化,误差单调下降但变化缓慢。由此可见,在迭代精度基本满足问题要求的前提下,一味增大种群规模以换取迭代精度所付出的计算代价巨大,因而是不可取的。本文的研究结果表明,将 GA 应用于电力系统综合负荷建模时,获得最佳的算法综合性能所对应的种群规模数以 100 左右为宜。基于此,本文推荐 $N=100$ 。

应当指出:本文所选取的反映 GA 性能的关键参数即为 t, g, e ;为了获得具有一般意义的结论,选取 10 个实测样本分别进行模型参数的优化辨识, t, g, e 均是 10 个样本所得相应结果的平均值;为了将 t, g, e 在同一坐标平面表示以方便分析,对它们进行了标准化使其在同等数量级。后文所给出的其他相应研究结果(图 2、图 3)均与此类似,不再说明。

1.2 交叉率和变异率

交叉率 p_c 决定了交叉算子的使用频率,其大小影响模式的生存和重组概率,经验值^[1]为 0.25~1.0;变异率 p_m 影响种群的多样性,控制了新基因导入种群的比例,其经验值^[1]为 0.005~0.1。

上述交叉率和变异率的经验值仍然是一个较大的取值范围。结合电力系统综合负荷建模应用实践,研究交叉、变异率对 GA 综合性能的影响,其结果如图 2、3 所示。

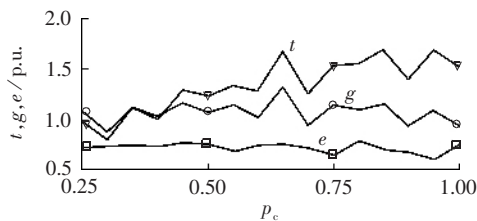


图 2 p_c 的变化对负荷建模的影响

Fig.2 Effect of crossover probability on load modeling

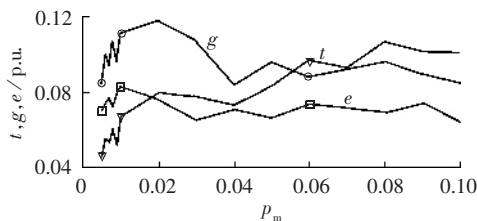


图 3 p_m 的变化对负荷建模的影响

Fig.3 Effect of mutation probability on load modeling

由图 2、3 分析可知,随着交叉率和变异率的增大,误差总体上呈单调下降趋势但变化不显著,收敛时间随之增长,而进化代数变化规律相对比较复杂。这是由于交叉率越大,群体中新模式产生的机会就会越多,在初始阶段探索领域甚至可以延伸到整个编码空间,但原有模式被破坏的可能性也随之增大,太低则可能导致搜索阻滞。与交叉算子相比,变异可以提高解群的多样性,但变异率太大则使得 GA 成为随机搜索,太小可能导致 GA 早熟,或者使 GA 陷入局部最优解的概率增大。同时应当看到,交叉率与变异率两者对性能的影响相互交织并具有相互牵制作用。因此,一方面选择合理的交叉、变异概率至关重要;另一方面,这 2 个参数的确定十分复杂,静态地确定其中的某个参数或者参数组合都不一定甚至很难获得理想的效果。因此,设计科学合理的策略使之在进化过程中动态地自适应调整变异率与交叉率,是解决此问题的有效途径。鉴于此,后文将对此进行进一步研究。

2 参数变量区间对负荷建模的影响

在电力系统负荷参数辨识研究中,模型参数的可辨识性^[6]和参数的稳定性问题一直倍受关注。鉴于此,进一步研究参数区间的改变对参数稳定性的影响。

对于感应电动机模型,以典型参数^[7]为参考,在不同样本的模型参数辨识中对参数区间合理调整以获得理想的参数稳定性。调整方法采用“经验试探法”。由于从参数的意义和优化结果看它们相互之间虽有联系但相对还是比较独立的,所以可先假定其他参数固定不变,研究单一参数的最佳选取值,最后得出综合的整体最优区间。

研究结果表明,表 1 所列为最佳的初始种群参数控制区间。

表 1 感应电动机参数区间

Tab.1 Induction motor parameter ranges p.u.						
参数值	r_s	X'	T_0'	n	T_j	a
最大值	0.5	4	5	5	10	5
最小值	0	0	0	0	0	0

表 1 的参数区间确定与各参数相应的物理意义^[7]及所采用的模型结构紧密相关,在实际建模工作中发现很多负荷辨识结果存在 r_s 和 X' 的值偏大,作者认为这是由于配电网线路未分开考虑而以串联的形

表 2 应用本文所列参数区间进行基于量测的负荷建模

Tab.2 Measurement-based load modeling by using parameter ranges listed											
数据	$E/\text{p.u.}$	G_c	t_c/s	$G/\text{p.u.}$	$B/\text{p.u.}$	$r_s/\text{p.u.}$	$X'/\text{p.u.}$	$T_0'/\text{p.u.}$	$n/\text{p.u.}$	$T_j/\text{p.u.}$	$a/\text{p.u.}$
Smdl1	0.00034	14	5	0.4889	1.0744	0.1390	0.3391	0.7561	1.8650	6.5199	1.8000
Smdl2	0.00036	21	6	0.4178	1.0367	0.1403	0.3409	0.7173	1.8664	6.5199	3.2300
Smdl5	0.00037	41	11	0.4836	1.0614	0.1401	0.3393	0.7617	1.8671	7.9594	1.4012
参数平均值				0.4635	1.0575	0.1398	0.3398	0.7450	1.8661	6.9998	2.1437
参数方差				0.0734	0.0301	0.0009	0.0013	0.0375	0.0015	1.1753	1.3599

注: E 为残差, G_c 为收敛代数, t_c 为计算时间;表 3 同。

由表 2 可见,本文所提出的参数区间是合理的,无论是单一参数还是总体效果,同名参数方差都相当小,辨识结果之参数稳定性很理想,这对于负荷建模的实用化具有重要意义。同时应当指出,机械参数 T_j 与 a 的辨识结果分散性相对较大,是由于这 2 个参数本身的不敏感性,而非算法原因^[8]。

3 p_c 和 p_m 的自适应控制策略

在 GA 的进化过程中,一般需要在初期提高模式的重组能力,加大编码空间探索的广度,以便找到最优解所在的子空间,跳出局部极值点或欺骗陷阱。在后期,希望提高模式的生存能力,以便在最优解子空间上逐步求精。从理论上讲,采用随着群体进化而变的交叉、变异概率,可以取得很好的效果。

因此,本文提出与群体进化程度指标相关的自适应调整交叉概率和变异概率的方法:在进化过程中根据解群的多样性程度自动地调整交叉概率和变异概率,使得 p_c 和 p_m 在不同代解群和不同个体的算子操作上采取不同的值,增强遗传算法可控制性,以期在解群进化过程中获得最佳交叉概率和变异概率。本文所设计的自适应策略模型如式(2)(3)所示。

$$p_c = \begin{cases} k_1 - k_2 / [k_3 + \exp(-k_7 \cdot \Delta')] & p_c > p_{c,\min} \\ p_{c,\min} & p_c \leq p_{c,\min} \end{cases} \quad (2)$$

$$p_m = \begin{cases} k_4 - k_5 / [k_6 + \exp(-k_8 \cdot \Delta')] & p_m > p_{m,\min} \\ p_{m,\min} & p_m \leq p_{m,\min} \end{cases} \quad (3)$$

其中, $p_{c,\min}$ 和 $p_{m,\min}$ 分别为交叉率和变异率设定最小值。 Δ' 表征解群内个体分布的多样性^[9]:设第 t 代解群由个体 $x_t^1, x_t^2, \dots, x_t^N$ 构成,适应度分别为 $f_t^1, f_t^2, \dots, f_t^N$, 个体总平均适应度 $\bar{f}_t = \frac{1}{N} \sum_{i=1}^N f_t^i$ (N 为解群规模);最优个体的适应度为 $f_{\max}$ 。将适应度大于 $\bar{f}_t$ 的

式计入了感应电动机定子回路。
具体参数区间由实际情况决定,其改变对 GA 辨识结果有很大的影响。区间过大,计算量增加使优化时间变长,且初始解不易分散到整个空间;区间过小,最优解一般只能在边界找到,结果不甚理想。
为检验上述所确定的模型参数控制区间对克服最终所获得的模型参数分散性的实际效果,依表 1 控制初始种群的参数变化范围,取 3 个属于同类负荷的样本分别独立地进行参数辨识,结果如表 2 所列。

个体的适应度平均得 $\bar{f}$, 定义 $f_{\max}$ 与 $\bar{f}$ 之间的差值 $\Delta' = f_{\max} - \bar{f}$ 。因此, Δ' 越小,表示种群中的各个体越相似,解群的多样性越差;相反, Δ' 越大,表示解群中的个体特性分散,解群的多样性就越大。 $k_1 \sim k_8$ 为交叉率和变异率的可调系数。

应强调指出,尽管有学者就 p_c, p_m 的自适应调整方法作过深入研究并设计了相应的调整策略^[9-10],但作者研究发现: $p_{c,\min}, p_{m,\min}$ 和 $k_1 \sim k_8$ 这 10 个可调参数的选取对算法性能的改善程度和速度有巨大的影响。曾有专家就此问题进行过深入研究并提出了参数调整的所谓“经验组合”^[1,11]。然而,经作者研究表明,对于不同的应用对象并无通用的取值准则,对于同样的应用类型(如负荷建模问题),是可以寻求到一组较佳组合的。作者通过大量的实测建模工作并结合具体情况考虑整个进化过程中交叉和变异概率的动态调整,深入分析后选取一组具有自适应特性的最佳控制参数组合: $p_{c,\min} = 0.4, p_{m,\min} = 0.06, k_1 = 1.3, k_2 = k_3 = k_4 = 1.0, k_5 = k_6 = 9, k_7 = 0.0001, k_8 = 0.0005$ 。

式(2)(3)中,由于 Δ' 始终非负,所以 p_c 和 p_m 是分别在区间 $[0.3, 0.8]$ 和 $[0, 0.1]$ 内随 Δ' 的增大而下降的变化曲线,可调系数 k_7, k_8 的大小影响 p_c, p_m 的变化速度。当 Δ' 变小,种群多样性降低而趋于收敛时, p_c 和 p_m 变大,增加交叉和变异的频率,加强种群的探测能力;当 Δ' 变大,种群多样性提升个体特征分散时, p_c 和 p_m 变小,降低交叉和变异的频率,使种群趋于收敛。通过上述分析可知, p_c 和 p_m 能在进化过程中根据种群多样性的增减,自适应地调整其大小。

为了检验作者所设计的自适应控制策略的性能,表 3 分别列出了采用本文提出的基于交叉概率和变异概率的适应性控制策略与固定参数的基本遗传算法(SGA)应用于综合负荷建模的结果比较。

表 3 SGA 与基于适应性交叉概率和变异概率的 AGA 的性能比较
Tab.3 Performance comparison between SGA and AGA based
on adaptive crossover and mutation probabilities

控制策略	数据 Smdl1			数据 Smdl2			数据 Smdl5		
	$E/p.u.$	G_c	t_c/s	$E/p.u.$	G_c	t_c/s	$E/p.u.$	G_c	t_c/s
$p_c=0.90, p_m=0.30$	0.000224	321	186	0.000400	136	61	0.00025	161	73
$p_c=0.75, p_m=0.10$	0.000216	201	57	0.000400	83	19	0.00025	1121	272
$p_c=0.55, p_m=0.10$	0.000230	261	67	0.000380	34	8	0.00027	221	46
$p_c=0.45, p_m=0.06$	0.000218	381	73	0.000396	71	11	0.00025	321	50
本文策略 AGA	0.000201	81	24	0.000396	102	22	0.00025	190	40

从表 3 可见,基于适应性交叉和变异概率的 AGA 比 SGA 的性能均有显著的提高——在同样精度要求下,其残差和收敛速度(收敛代数/计算时间)都远优于 SGA。而且应当强调指出:尽管对于某些样本辨识而言,固定的 p_c, p_m 也能获得很好的辨识结果,但是这种固定组合是经大量实验试探获得的,而且不具有对其他样本的普遍适应性。因此,自适应的 p_c, p_m 控制策略具有固定方式不可比拟的优势。

4 结语

本文以基于实测的电力系统综合负荷建模为应用对象,深入研究了遗传算法中种群规模、交叉、变异概率及其控制策略,以及初始种群参数区间等遗传算法关键操作参数对算法性能的影响规律。研究结果表明:种群规模 $N=100$ 为最宜,交叉与变异概率应当采用自适应策略;参数初始区间应当根据具体情况予以调整。作者成功地将上述规则应用于电力系统综合负荷建模,检验了工作的有效性。本文工作对 GA 应用于其他工程优化问题具有参考与借鉴意义。

参考文献:

[1] 李敏强,寇纪淞. 遗传算法的基本理论与应用[M]. 北京:科技出版社,2003.

[2] 玄光男,程润伟. 遗传算法与工程优化[M]. 于歆杰,周根贵,译. 北京:清华大学出版社,2004.

[3] 张勇军,任震,唐卓尧,等. 电压无功优化的强多样性遗传算法[J]. 电力自动化设备,2003,23(1):18-24.

ZHANG Yong-jun,REN Zhen,TANG Zhuo-yao,et al. Diversity enhancing genetic algorithm for voltage and reactive power optimization[J]. **Electric Power Automation Equipment**,2003,23(1):18-24.

[4] 李欣然,贺仁睦,周文,等. 综合负荷感应电动机模型的改进及其描述能力[J]. 电力系统自动化,1999,23(9):23-27.

LI Xin-ran,HE Ren-mu,ZHOU Wen,et al. The generalized induction motor model and its description ability to synthetic loads of electric power system[J]. **Automation of Electric Power Systems**,1999,23(9):23-27.

[5] 李欣然,刘艳阳,陈辉华,等. 遗传算法与传统优化方法应用于电力负荷建模的比较研究[J]. 湖南大学学报,2005,32(2):29-32.

LI Xin-ran,LIU Yan-yang,CHEN Hui-hua,et al. Comparative study of genetic algorithm and traditional optimization in power load modeling[J]. **Journal of Hunan University**,2005,32(2):29-32.

[6] 鞠平,赵夏阳,李东辉. 电力负荷模型可辨识性分析方法[J]. 电力系统自动化,1999,23(19):29-33.

JU Ping,ZHAO Xia-yang,LI Dong-hui. Approaches to identifiability analysis of electric load models[J]. **Automation of Electric Power Systems**,1999,23(19):29-33.

[7] 倪以信,陈寿孙,张宝霖. 动态电力系统的理论和分析[M]. 北京:清华大学出版社,2002.

[8] 张红斌. 电力系统负荷建模模型结构与参数辨识的研究[D]. 保定:华北电力大学,2003.

ZHANG Hong-bin. The power system load model structure and the parameter recognizes[D]. Baoding:North China Electric Power University,2003.

[9] 吴浩扬,朱长纯,常炳国,等. 基于种群过早收敛程度定量分析的改进自适应遗传算法[J]. 西安交通大学学报,1999,33(11):27-30.

WU Hao-yang,ZHU Chang-chun,CHANG Bing-guo,et al. Adaptive genetic algorithm to improve group premature convergence[J]. **Journal of Xi'an Jiaotong University**,1999,33(11):27-30.

[10] 盛兆俊,刘瀚. 基于改进遗传算法的无功综合优化[J]. 电力自动化设备,2004,24(4):27-29.

SHENG Zhao-jun,LIU Han. Reactive power optimization of integrative power system based on improved genetic algorithm[J]. **Electric Power Automation Equipment**,2004,24(4):27-29.

[11] de JONG K A. An analysis of the behavior of a class of genetic adaptive systems[D]. East Lansing,USA: University of Michigan,1975.

(责任编辑:李育燕)

作者简介:

金 群(1981-),女,上海奉贤人,硕士研究生,主要研究方向为电力系统动态负荷建模(E-mail:jinqun1981@126.com);

李欣然(1957-),男,湖南涟源人,教授,博士研究生导师,主要从事电力系统分析与控制及负荷建模的教学和研究工作(E-mail:lixr1013@yahoo.com.cn)。

GA parameter setting and its application in load modeling

JIN Qun,LI Xin-ran

(Hunan University,Changsha 410082,China)

Abstract: Combining with the application of GA(Genetic Algorithm) in power system load modeling based on measured field data,the operation mechanism of genetic algorithm is discussed and the search ability of different genetic operators is analyzed. The population variety is the key factor determining GA performance and the relative restraints between population variety and algorithm parameter are obtained. Rules of genetic parameter setting are provided from theoretical analysis,and the population scale,crossover probability,mutation probability and their control strategies are studied,as well as the influence of key operation parameters,such as initial parameter ranges,on the performance of GA. Reasonable population scale,initial parameter ranges are put forward,and the self-adaptive modulation strategies for crossover and mutation probabilities,which are relative to innovation degree,are proposed. The result of study indicates that appropriate parameter combination is the key to excavate the potential of GA,which improves its operational efficiency,avoids precocity and reduces the dispersiveness of model parameters.

This project is supported by the National University Primary Teacher Foundation of China ([2002]65) and Hunan Province Department of Education Project([2001]197).

Key words: electric power system; load modeling; genetic algorithm; control parameter; parameter range
