

基于复杂系统演化优化的实码多目标嵌套加速遗传算法

付 强^{1,2}, 王 凯¹, 任守德¹

(1. 东北农业大学 水利与建筑学院, 哈尔滨 150030; 2. 黑龙江省高校节水农业重点实验室, 哈尔滨 150030)

摘 要 复杂性科学研究已经成为系统科学领域研究的重要方向, 然而由于复杂系统的多层次、非线性等特点, 加之含有大量随机不确定因素, 给系统的演化仿真工作带来巨大的困难, 提出基于实数编码的多目标嵌套加速遗传算法 (RMONAGA) 来对复杂系统进行演化仿真, 利用复杂适应系统自身特点与遗传算法特点的相似性, 采用嵌套、多目标等方法较好的解决了复杂系统中不同层次交互、不同部门间适应性矛盾等问题. 将该方法应用于三江平原区域农业水土资源复杂适应性优化配置演化研究中, 验证了该方法的科学合理性, 以期对 CAS 理论在各领域的建模仿真工作提供一定的帮助.

关键词 遗传算法; 复杂适应系统; 演化仿真

Real coded multi-objective nested accelerating genetic algorithm based on complex system evolutionary optimization

FU Qiang^{1,2}, WANG Kai¹, REN Shou-de¹

(1. Collage of Water Conservancy and Architecture, Northeast Agricultural University, Harbin 150030, China; 2. Key Laboratory for Water-saving Agriculture of Universities in Heilongjiang Province, Harbin 150030, China)

Abstract The research of complex system science has become an very essential part of the system science. However, due to it's multi-level, nonlinear, and coupled with a large number of random uncertainties in the evolution of simulation work to the system caused immense difficulties. On this background, the real coded multi-objective nested accelerating genetic algorithm (RMONAGA) has been proposed to evolve and simulate the complex system. Based on the similarity between the features of system and characteristics of this algorithm, the problem of interaction of different levels and the adaptive conflict between different sectors has been solved by using the nested and multi-objective method. The method is applied to do the research of regional agricultural soil and water resources complex adaptive system optimization of Sanjiang Plain. The result shows that this method is scientifically rational, in order to provide some help in the field of modeling and simulation when CAS theory have been applied.

Keywords genetic algorithm; complex adaptive system; evolution and simulation

1 绪言

随着人类认识的不断深入与发展, 人们发现某些系统总是无法用标准化的分析方法来理解和解决, 它们是由许多相互作用的实体组成的, 相互之间有着极为复杂的联系, 这些系统被称为复杂系统. 近年来, 复杂性科学的研究已经成为系统科学领域研究的重要方向. 复杂适应系统 (complex adaptive system, CAS) 作为一类具有代表性的复杂系统, 备受学术界关注, 被广泛应用于自然科学、经济、国防军事、社会科学、人文科学等各个领域^[1-5].

然而, 由于复杂适应系统的内部结构和关系极为复杂, 涉及到大量非线性参数优化和多目标等问题, 导致

收稿日期: 2010-08-24

资助项目: 国家自然科学基金 (51179032); 教育部高等学校博士学科点专项科研基金 (20092325110014); 哈尔滨市科技创新人才专项资金 (RC2010XK002013); 黑龙江省普通高校新世纪优秀人才培养计划 (1155-NCET-004); 教育部新世纪优秀人才支持计划; 国家水利公益性行业科研专项项目 (200901075)

作者简介: 付强 (1973-), 男, 辽宁锦州人, 教授, 博士生导师, 研究方向: 农业水土资源系统分析, 节水灌溉及农业系统工程建模与优化技术.

复杂适应系统的优化演化和仿真分析存在很大的困难. 研究人员花费了大量的精力对复杂适应系统的内部关系进行梳理, 但就其演化仿真仍然遇到很大困难. 许多学者也针对这样的问题进行了研究, 如赵建世^[6]建立了基于 CAS 理论的水资源优化配置整体模型, 采用嵌套遗传算法求解非线性模型, 采用情景分析法和交互的切比雪夫方法解决多目标问题; 牛文娟等^[7]应用 CAS 理论建立南水北调东线工程水资源配置的 3 层仿真模型; 邓宏钟、谭跃进^[8]提出基于复杂适应系统的复合多层次遗传算法分别对模型层、Agent 层和参数进行演化优化. 但是由于这些方法涉及的理论范围较广, 演化过程复杂, 对计算机的计算能力要求较高, 一般的学者不易掌握, 即使能够进行演化分析所耗费的时间也较长.

针对上述问题, 作者将传统的遗传算法进行了改进和耦合, 提出了一种基于实数编码的多目标嵌套加速遗传算法 (RMONAGA) 来对复杂适应系统进行优化演化和仿真分析. 该方法所涉及的理论基础均与遗传算法一脉相承, 运行效率较高, 便于理解和实践应用, 以期对 CAS 理论在各领域中的建模仿真工作提供一定的帮助.

2 复杂系统演化优化的关键问题

基于实数编码的多目标嵌套加速遗传算法 (RMONAGA) 是在解决复杂系统优化演化过程中遇到的实际问题而产生的, 对于复杂适应系统理论建立的模型来讲, 不同部门有着不同的适应性, 各层次间流的传递也有所不同. 故在不同的部门和层次中采用了不同的遗传算法模拟求解, 采用嵌套的方式组合在一起, 整体上便形成了该算法. 下面针对优化演化过程中的几个关键问题对算法进行介绍.

2.1 不同主体及部门间的适应性矛盾

一般情况下, 复杂系统由于时空分布上的差异会导致其具有很强的层次性, 形成了一个由上而下, 由点到面的多层次多功能空间网络体系, 其中会包含多个子系统, 如社会、经济、生态等子系统. 同时, 系统的功能也具有与其结构层次相类似的层次特点, 这就使得系统内的不同主体及部门间有着不同的适应性, 一个主体或部门适应性的改善, 其代价往往是其它适应性的恶化. 如何权衡不同主体和部门间适应性的问题成为系统能否达到整体最优状态的关键.

这里采用基于排序选择法的多目标遗传算法 (MOGA) 来解决上述问题. 同标准遗传算法相比该方法的关键在于适应度函数的构造, 关于多目标遗传算法的其它步骤这里不再赘述, 读者可以参看相关文献^[9-10]进行学习. 基于排序计算适应度的方法只取决于多目标的本身, 故可采用将种群中所有个体对不同目标函数的优劣进行排序, 从而来计算总的适应度.

用 $Z(i)(i = 1, 2, \dots, n)$ 表示目标函数, n 为目标个数, 对于每一个目标 i , 所有个体都会依据对该目标的函数值优劣生成一个可行解的排序序列 X_i . 对每一个目标都排序后, 可以得到个体对全部目标函数的总体表现. 根据个体的排序计算其适应度:

$$F(X_j) = \begin{cases} (N - Y_i(X_j))^2, Y_i(X_j) > 1, \\ kN^2, Y_i(X_j) = 1, \end{cases} \quad i = 1, 2, \dots, n \quad (1)$$

$$F(X_j) = \sum_{i=1}^n F_i(X_j), \quad j = 1, 2, \dots, n \quad (2)$$

式中: n 为目标函数总数; N 为个体总数; X_j 为种群的第 j 个个体; Y_i 为其在种群所有个体中对目标 i 的优劣排序后所得的序号; $F_i(X_j)$ 表示 X_j 对目标 i 所得的适应度; $\sum_{i=1}^n F_i(X_j)$ 为 X_j 对全部目标所得的综合适应度; k 为 $(1, 2)$ 区间的常数, 用于加大个体的函数值表现最优时的适应度; 由上式可以看出, 对于总体表现较优的个体能得到更大的适应度, 获得更多的参与进化的机会, 从而被保留下来.

例如, 对于农业部门来讲, 产量与产值之间存在一定矛盾, 如何优化作物种植的面积以权衡产量与产值得到最大的农业效益, 可以将各方案所得产量与产值进行排序, 根据式 (1)、(2) 计算 $F_i(X_j)$, 作为评价该方案农业综合效益的适应度, 再进行遗传演化得到农业部门最大综合效益下的方案组合.

2.2 系统各层级间“流”的传递与交互

主体与主体, 主体与环境间的相互作用, 使得个体的变化成为整个系统变化的基础, 而在个体与环境之间存在有物质、能量和信息流, 这些流的渠道是否通畅、周转迅速到什么程度, 都直接影响系统的演化过程. 上层主体与下层主体间通过不断反复的交互, 实现信息在不同层次间主体的流动, 最终达到整体的综合效益最大.

根据模型不同层次的特点用相应的遗传算法模拟演化,整体上便形成了嵌套遗传算法.所谓嵌套遗传算法(NGA),就是在一个单独的遗传算法内部嵌入另一个遗传算法,用以优化外层遗传算法的一个或多个参数,最终得到外层算法需要优化的参数.具体的步骤可以结合实际问题来完成.这种算法的基本流程如图 1 所示.

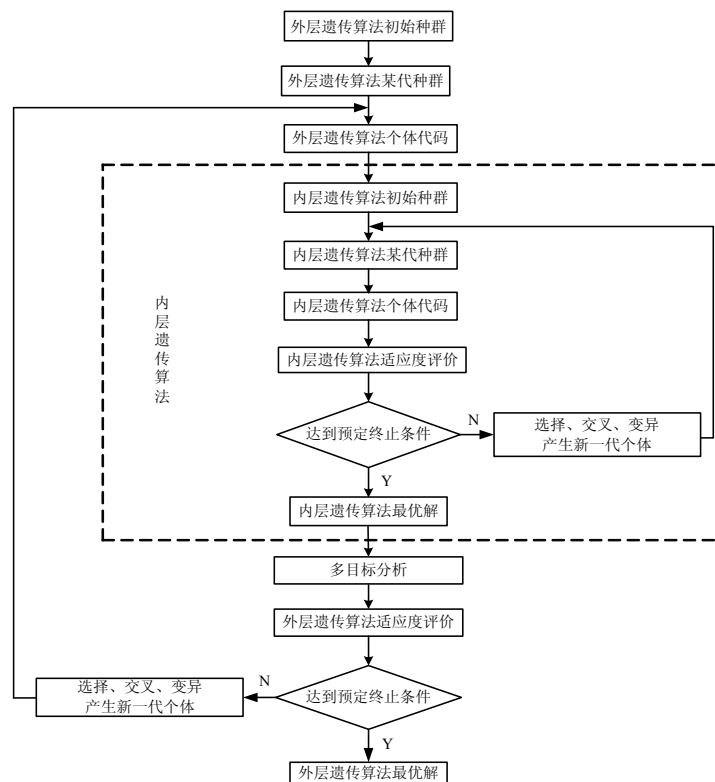


图 1 嵌套遗传算法流程图

对于含有多个子系统的复杂系统来讲,可将经济、社会、生态等子系统用单独的遗传算法求解,其结果作为外层遗传算法的输入,用以评价系统整体的状态,再通过外层算法的优化,各层次间不断的交互,获得系统的最优状态.

2.3 提高算法的运行效率

由于复杂系统高度的开放性和层次性,使得系统内部、系统与环境之间存在着极为复杂的联系,加之含有大量不确定随机因素与非线性优化问题,导致与多目标及嵌套遗传算法相结合后运行效率较低,采用如下方法可在一定程度上提高算法整体运行效率.

1) 实数编码

标准遗传算法(SGA)中采用的二进制编码在进行各种遗传操作时需要频繁的编码解码,同时编码过长,占用较多的计算机存储空间,还会由于优化变量维数过高增加算法的计算量,使得遗传算法的收敛速度降低,甚至无法接受.这里采用实数编码的方法对优化变量进行编码,可以采用如下线性变换:

$$x(j) = a(j) + y(j)(b(j) - a(j)), \quad j = 1, 2, \dots, p \quad (3)$$

上式把初始变量区间 $[a(j), b(j)]$ 上的第 j 个待优化变量 $x(j)$ 对应到 $[0, 1]$ 区间上的实数 $y(j)$, $y(j)$ 即为遗传基因.此时,优化问题所有的变量对应的基因顺次连在一起构成问题解的编码形式 $(y(1), y(2), \dots, y(p))$,称之为染色体.经编码,所有优化变量的取值范围均变为 $[0, 1]$ 区间,于是可以直接对各优化变量的基因进行选择、交叉、编译等遗传过程的各种操作.

2) 加速处理

根据对 GA 的选择、杂交、变异这三种算子的寻优性能的分析 and 大量的数值实验与实际应用,这里采用上一次演化迭代所产生的优秀个体这一子群体所对应的变量变化区间,作为变量新的初始变化区间,而后进入新一轮演化迭代.如此加速循环,优秀个体的变化区间将逐步调整和收缩,与最优点的距离将越来越近,直至最优个体的目标函数值小于某一设定值或算法运行达到预定加速(循环)次数,结束整个算法的运行,并把当前群体中最佳个体或优秀个体的平均值指定为最后的结果.

3) 优秀个体保存

最优保存策略可以采用“移民”操作. 具体步骤如下: 1) 首先计算当前种群中个体适应度, 找出适应度最大和最小的个体; 2) 对种群中除去适应度最大的 n 个个体之外的其它个体进行交叉变异操作, 然后计算新种群的适应度值, 找出 n 个适应度最小的个体; 3) 用预先保留的 n 个个体替换遗传操作后的适应度小的 n 个个体, 并形成新种群. 一般设置每次保留适应度值最大的 5 个个体, 也可以根据种群规模设定, 一般为种群个体总数的 5%–15%.

3 RMONAGA 整体框架

结合复杂适应系统思想建立的优化配置模型, 最大的特点就是高维度和非线性, 这也是模型优化仿真的难点. 模型的构建以主体即模型中的最终用户为基本单位, 这些用户具有不同的资源利用方式. 这些行为方式间的相互作用“涌现”出了更高层次的部门级主体资源利用的行为特征. 部门级资源行为之间的相互作用“涌现”出了更高层次的政府级主体资源配置的行为特征, 政府级主体追求的是整体综合效益的最大, 用以权衡多个部门的利益. 针对这样的特点以及上文的论述, 对不同部门采用不同的遗传算法, 采用嵌套的方式结合到一起, 整体上形成了基于实数编码的多目标嵌套加速遗传算法 — RMONAGA, 如图 2 所示.

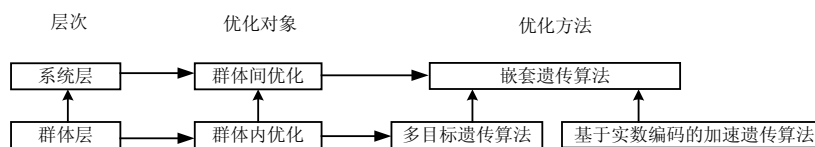


图 2 遗传算法在模型不同层次中的应用

正是由于模型基于主体和层次的框架, 使得模型的不同层次可以和相应的遗传算法相结合, 整体上形成了嵌套遗传算法的求解方法. 对部门级主体采用独立的遗传算法求解, 并将其优化的结果, 即部门级主体的输出, 作为政府级主体遗传算法的输入, 为模型最高层次的遗传算法的个体适应度评价提供支持, 如图 3. 通过这样的分解, 可以大大降低整体模型求解的复杂程度, 缩短求解的时间并提高求解精度.

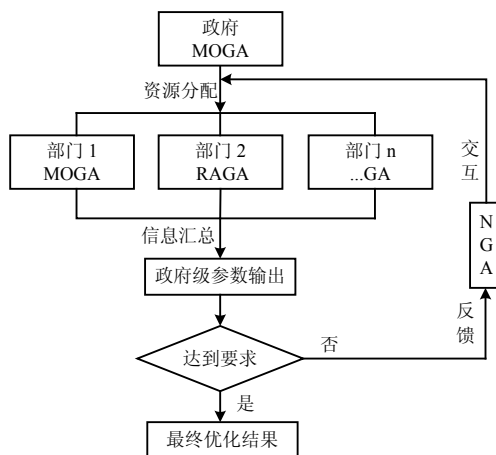


图 3 RMONAGA 模拟仿真框架图

通过上述分析, 遗传算法及其改进算法可以有效地解决复杂适应系统优化配置模型的演化仿真问题. 而这都源于复杂适应系统自身特点与遗传算法特点的相似性. 在应用遗传算法求解问题时, 在确定了编码方案、适应度函数及遗传算子以后, 算法将利用演化过程中获得的信息自行组织搜索. 遗传算法的自组织、自适应特征同时也赋予了它具有能根据环境的变化自动发现环境的特性和规律的能力, 同时很好的模拟和解决了复杂适应系统中主体的智能性, 包括自组织、自适应和自学习性. 遗传算法的内在并行性, 使它可以同时搜索解空间内的多个区域, 并相互交流信息. 通过深入地思考, 我们可以发现这种类似性完全源于基本思想和基本理论的一致性 [8,11].

4 参数设置

ROMNAGA 的关键控制参数包括内层与外层算法的群体规模 N 和 NN 、内层与外层算法的加速次数

Z 和 ZZ 、优秀个体数目 d 等, 必须对它们进行适当的设置才能达到 ROMNAGA 运行的最优性能。

1) 内层算法. N 与 Z 是相互平衡的关系, 而这之间的关系直接影响局部多样性, 间接影响全局最优解的搜索的实现. N 越大内层算法多样性越好, 可向外层算法提供较好的参数输出, 但由于内层算法迭代次数较多可能会使运算时间加长; Z 越小内层算法运行时间越短, 但有可能出现早熟收敛. 根据经验, 常用的搭配有 (200,10), (300,20).

2) 外层算法. NN 通过影响全局多样性, 可直接对最终的结果造成影响, 最小值为 50, 若 $NN = 100$ 优化时间明显变长, 但优化效果并无明显的增强. 因此最佳取值范围 $NN = [50, 100)$; 当 $ZZ = 10$ 时, 运行速度快, 但会出现早熟收敛, $ZZ = 20$, 程序会在第 18 次左右得到最优结果, 故最佳取值范围为 $ZZ = (10, 20]$; d 越大, 优秀个体包围、接近最优点的机会就越大, 但收敛速度越慢, 通过实验 d 的最佳取值范围为 [5,10).

5 实例研究

将基于实数编码的多目标加速遗传算法应用于区域农业水土资源复杂适应性配置仿真当中. 针对不同的地下水开发方案 (见表 1), 对当地种植结构进行优化仿真配置.

表 1 不同方案地下水开采量 ($10^8 m^3$)

方案编号	1	2	3	4	5	6	7	8	9	10
地下水开采量	23.80	23.50	23.00	22.47	21.64	21.19	20.94	20.36	19.90	19.47

针对该模型的特点, 在政府级遗传算法当中嵌套了两个部门级遗传算法: 农业部门及生态部门遗传算法. 农业部门具有多目标的特点, 采用多目标遗传算法对该部门进行演化仿真, 采用实码加速遗传算法对生态部门进行演化仿真. 仿真结果反馈给政府级主体, 为政府级模型的遗传算法提供适应度评价参数, 政府级主体将部门反馈的信息进行汇总, 再根据区域综合效益最大为目标进行资源重新分配, 通过这样不断反复的交互, 实现信息在不同层次间主体的流动, 最终达到区域综合效益最大.

以方案一为例, 具体演化过程如图 4. 其中, NN 、 N 分别是外层与内层算法初始父代规模; Pc 、 Pm 为交叉及变异概率; W_s 为可用地下水量; n 为优化变量个数; C 为加速次数; T 为交互次数.

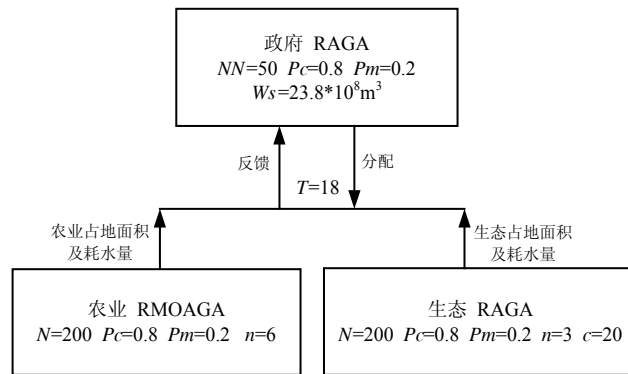


图 4 演化过程细节图

仿真结果如图 5、6 所示. 结果真实地反映了该地区不同的地下水开采强度下的种植结构方案及其与粮食产量、产值、生态面积及高效利用土地面积的关系, 得到了水资源已经成为制约当地经济社会及生态发展的瓶颈因素的结论, 验证了该方法的科学合理性.

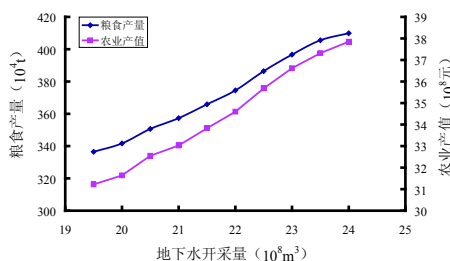


图 5 地下水开采量与粮食产量农业产值关系图

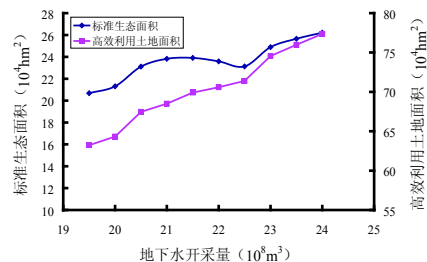


图 6 地下水开采量与标准生态面积和高效利用土地面积关系图

6 结论与展望

复杂适应系统理论为探究复杂系统提供了一条新的途径, 但就系统建模优化仿真工作仍然存在很大的困难, 如系统各层级间流的传递及交互的问题、不同部门适应性的矛盾问题、大量不确定随机因素与非线性优化造成的运行效率过低的问题等. 针对上述问题, 本文提出了基于实数编码的多目标嵌套加速遗传算法 (RMONAGA) 来对复杂系统进行演化仿真. 分别采用嵌套、多目标加速等方式较好地解决了上述存在的问题. 通过实例研究说明 RMONAGA 具有较好的收敛性和较强的可操作性及稳健的全局优化能力, 验证了该方法的科学性, 可以将其应用于复杂系统的仿真演化研究中.

参考文献

- [1] 蔡琳, 薛惠锋, 寇晓东. 基于 CAS 的城市空间演化多主体模型方法研究 [J]. 计算机仿真, 2007, 24(4): 145–148.
Cai L, Xue H F, Kou X D. Research on urban spatial evolution multi-agent modeling method based on CAS[J]. Computer Simulation, 2007, 24(4): 145–148.
- [2] 迟妍, 邓宏钟, 谭跃进, 等. 军事复杂适应系统理论基本框架研究 [J]. 军事运筹与系统工程, 2004, 18(2): 13–17.
Chi Y, Deng H Z, Tan Y J, et al. Basic framework study of military complex adaptive system theory[J]. Military Operations Research and Systems Engineering, 2004, 18(2): 13–17.
- [3] 毕贵红, 王华. 固体废物管理系统演化与调控的复杂适应系统理论分析 [J]. 工业技术经济, 2009, 28(2): 96–102.
Bi G H, Wang H. Model of solid waste management system evolution and adjustment based on complex adaptive system[J]. Industrial Technology and Economy, 2009, 28(2): 96–102.
- [4] 杨顺顺, 栾胜基, 王颖, 等. 基于 CAS 理论的循环经济适宜性评价与实证研究 [J]. 北京大学学报: 自然科学版, 2008, 44(4): 611–617.
Yang S S, Luan S J, Wang Y, et al. The feasibility analysis and case study for developing circular economy based on complex adaptive system theory[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2008, 44(4): 611–617.
- [5] 尹春华, 方福康. 复杂适应系统 Internet 中的网络流量仿真研究 [J]. 系统工程学报, 2005, 20(2): 139–142.
Yi C H, Fang F K. Study on network traffic simulation of a complex adaptive system-Internet[J]. Journal of Systems Engineering, 2005, 20(2): 139–142.
- [6] 赵建世. 基于复杂适应系统理论的水资源优化配置整体模型研究 [D]. 北京: 清华大学, 2003.
Zhao J S. Study on holistic model for optimization of water resources allocation based on complex adaptive system theory[D]. Beijing: Tsinghua University, 2003.
- [7] 牛文娟, 王慧敏. 基于 CAS 理论的南水北调东线水资源优化配置模型 [J]. 河海大学学报: 自然科学版, 2007, 35(4): 384–387.
Niu W J, Wang H M. CAS-based water resources optimal allocation model for east route of South-to-North Water Transfer Project[J]. Journal of Hohai University: Natural Sciences, 2007, 35(4): 384–387.
- [8] 邓宏钟, 谭跃进. 基于复杂适应系统的复合多层次遗传算法 [J]. 计算机工程与应用, 2002, 1: 22–24.
Deng H Z, Tan Y J. Composite multilayer genetic algorithms based on complex adaptive system[J]. Computer Engineering and Applications, 2002, 1: 22–24.
- [9] 刘攀, 郭生练, 李玮, 等. 用多目标遗传算法优化设计水库分期汛限水位 [J]. 系统工程理论与实践, 2007, 27(4): 81–90.
Liu P, Guo S L, Li W, et al. Optimal design of reservoir seasonal flood control level by using multi-objective genetic algorithm[J]. Systems Engineering — Theory & Practice, 2007, 27(4): 81–90.
- [10] 孙月峰, 张胜红, 王晓玲, 等. 基于混合遗传算法的区域大系统多目标水资源优化配置模型 [J]. 系统工程理论与实践, 2009, 29(1): 139–144.
Sun Y F, Zhang S H, Wang X L, et al. Multi-objective optimization of regional water resources based on mixed genetic algorithm[J]. Systems Engineering — Theory & Practice, 2009, 29(1): 139–144.
- [11] 牛文娟, 王慧敏. 基于遗传算法的水资源复杂适应配置研究 [J]. 水利经济, 2007, 25(5): 8–10.
Niu W J, Wang H M. Study on water resources complex adaptive allocation based on genetic algorithm[J]. Journal of Economics of Water Resources, 2007, 25(5): 8–10.