

文章编号:1001-9081(2006)08-1947-03

求解动态组播路由问题的混合优化遗传算法

陈云亮¹, 杨捷², 康立山¹

(1. 中国地质大学 计算机学院, 湖北 武汉 430074;

2. 华中科技大学 图像识别与人工智能研究所, 湖北 武汉 430074)

(cyl_king@hotmail.com)

摘 要:分析了具有网络时延和时延抖动限制的动态组播路由问题的数学模型。在此模型的基础上提出了一种基因库(GP)与传统遗传算法(GA)混合的优化算法 GP-GA。该算法利用基因库保存进化过程中得到的解路径以指导后继进化过程,同时改进了交叉和变异算子来加快算法的收敛速度。考虑到问题可能陷入的局部最优情况,又构造了基于“保留和不保留”的进化控制策略来增强寻优能力,很大程度上避免了算法“早熟”现象的发生。大量的仿真实验表明:GP-GA 算法相对现有的遗传算法求得最优解的概率更高,相对于动态的组播环境也有很好的代价性能。

关键词:Steiner 树;动态组播路由;基因库;遗传算法;路由优化

中图分类号:TP183 **文献标识码:**A

Hybridized optimization genetic algorithm for multicast routing problem

CHEN Yun-liang¹, YANG Jie², KANG Li-shan¹

(1. School of Computer, China University of Geosciences, Wuhan Hubei 430074, China;

2. Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan Hubei 430074, China)

Abstract: The mathematic model of dynamic multicast routing with Nodes delay and delay variation constraints was analyzed. Based on the model, an optimization algorithm called GP-GA was proposed by hybridizing Gene-Pool (GP) with traditional Genetic Algorithm (GA). This method made use of the gene-pool to save the solutions during the process so as to direct the remaining evolution. In the mean time, the crossover and mutation operator were improved to accelerate the convergence speed. Considering that the problem may be trapped by local optimization easily, the evolution strategy based-on "reserved and non-reserved" was also constructed to enhance the ability of finding optimal solution and decrease the probability of "premature" phenomena commendably. A great number of simulations demonstrate that the probability of GP-GA converging optimal solutions is higher than general GA, and the algorithm is also effective for being adjusted to the dynamic multicast routing.

Key words: Steiner tree; dynamic multicast routing; Gene-Pool(GP); Genetic Algorithm(GA); routing optimization

组播路由就是要寻找一棵费用最小的连接源节点和一组目的节点的组播树,这个问题在数学上归结为最小 Steiner 树问题。同时考虑到实际情况中目的节点动态增加和删除,构造动态组播路由也成为研究的内容。由于该问题已被证明为是 NP-Complete 问题,目前大都采用启发式算法。

文献[1]采用 Hopfield 神经网络给出了计算满足特定时延和时延抖动要求的 Steiner 树;文献[2]提出了运用蚁群算法进行并行组播寻路的算法模型;文献[3]介绍了通过接种疫苗和免疫选择两个步骤构造免疫算子,并将免疫算法应用于组播路由选择的方法;文献[4]提出的算法利用禁忌搜索方法中记忆功能和禁忌规则等特点来构造时延约束最小代价组播树;文献[5]利用模拟退火算法的思路,构造了降温(退火)进度表来模拟对应路由树的网络延迟和费用等,从而逼近问题的最优解。在遗传算法解决该问题的过程中,研究人员设计了各种遗传算子,文献[6]的原理是利用遗传算法对网络中可能构成的 Steiner 节点组合进行搜索,采用传统的二进制编码。文献[7~11]利用组播树上从已给的源结点 s 到

目标节点 t 的可能路径 (s, t) 作为染色体上的基因片,从而进行组合得到整条染色体的编码方案。

本文在已有研究的基础上,提出了一种基于基因库(Gene-Pool, GP)解决 Steiner 问题的遗传算法模型。在遗传算子勘探解空间的过程中,基因库可获得并保存好的基因块而不致流失。同时改进了传统 GA 算法解决该问题时采用的交叉和变异算子,并对控制选择策略进行了修正,很好地避免了“早熟”现象的发生。实验表明,该算法模型求得最优解的概率相比已有的 GA 算法有了较大的提高。

1 求解 Steiner 树问题的 GP-GA 算法

1.1 染色体和基因库的初始化

本文中提出的染色体由两部分路径组成,即(源节点 s 到某中继节点 k 的路径,中继节点 k 到某目的节点 t 的路径)。本算法在初始化群体采用了以下的策略:随机生成从源节点 s 到中继节点 k 的路径,或取源节点 s 到中继节点 k 的最短路径作为染色体的前半部;之后随机选择一个与该中继节点 k 相

收稿日期:2006-03-01;修订日期:2006-04-18 基金项目:中国地质大学(武汉)优秀青年教师资助计划资助项目(CUCQN144)

作者简介:陈云亮(1979-),男,江苏如皋人,讲师,主要研究方向:高级网络管理、智能计算; 杨捷(1981-),男,福建福州人,硕士研究生,主要研究方向:智能计算、遗传算法; 康立山(1934-),男,湖南衡山人,教授,博士生导师,主要研究方向:高性能计算机、智能计算。

邻的目的节点 t 做连接,从而组成染色体。算法中提出的基因库被设计为一个采用了数组指针的十字链表,其中数组的元素用来指示每个中继节点,其对应连接的十字链表用来保存进化过程中出现的从源节点 s 到中继节点 k 的路径,它可以来源于最优甚至是最差解。

1.2 交叉算子

传统遗传交叉算子有单点交叉,多点交叉和线性重组。在实验中我们采用了 Disjoin 的交叉算子,因为它基本保存了父代已形成的路径,因而加快了收敛度。Disjoin 算子的形式描述具体算法见文献[12]。将 Disjoin 交叉算子与文献[7,8,12~15]中传统的 One-point 交叉算子进行比较,我们发现前者的收敛速度更快。

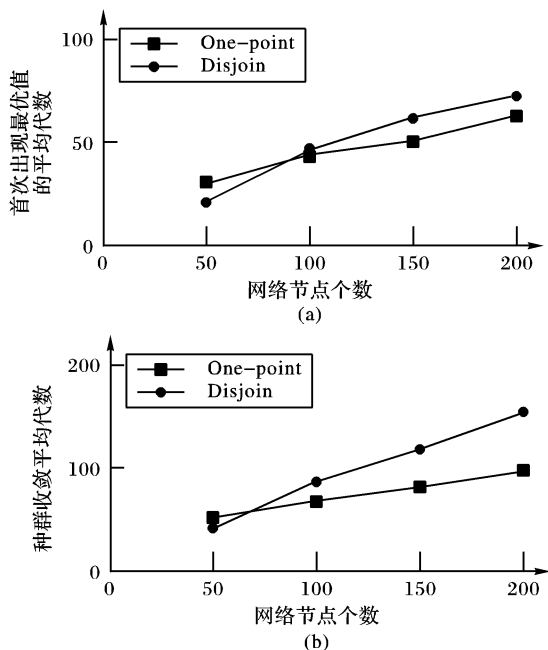


图1 Disjoin 算子与 One-point 算子的比较结果

由此我们得出结论:在相同的情况下,Disjoin 算子的进化速度明显高于传统的 One-point 交叉算子,尽管在网络节点个数较少($=50$)的情况下,Disjoin 算子会略逊于 One-point 交叉算子。但是随着个数的增加,越来越多的中继节点和边被加入到染色体中。而一旦产生交叉,One-point 算子得到的子代相对 Disjoin 算子需要作更多的缝合(增加新边)和删除(避免环)操作,这样 Disjoin 算子的优越性就被体现出来了。但同时我们也注意到,一旦进化过程中出现一些超常的个体,若按照传统的种群选择策略,这些异常的个体将由于竞争力太突出(即适应度较高)而控制了剩下的进化过程。从图 1 不难看出,在最优个体出现的代数的后 50~60 代,算法基本就趋近收敛了,这严重影响了算法的全局优化性能,最后获得的很可能只是某个局部最优解。为解决这种典型的遗传欺骗问题,我们同时改进了进化策略,详见 1.4。

1.3 变异算子

除采用传统的变异算子(即利用随机生成的一条从源节点到某目的节点的路径代替染色体中相同终点的路径外),还设计了如下两种变异算子:

Mutation 1: 依变异概率 P_{m1} 随机选择一条染色体中的路径,用基因库中具有相同终点的路径替代它。

Mutation 2: 依变异概率 P_{m2} 随机选择一个目的节点 t , 剪断其与中继节点 ki 的边,将它连接到与之相邻的另一个中继节点 kj 上。

显然 Mutation 1 在操作的时候参考了基因库中已保存的

路径,这样收敛的速度得到了提高;而 Mutation 2 则用于搜索局部范围内的优秀个体。

1.4 进化策略

我们发现很多文献均采用了传统的赌轮法(又叫 Monte Carlo 选择法)作为选择染色体的策略。但这样做容易过早收敛且在进化后期搜索效率低,最终搜索到的结果不是全局最优解的概率很大。文献[15]中在双种群的基础上引进了模拟退火更新策略来改进对群体的选择,得到了很好的效果。我们由此得到启发,提出了“Reserved and Non-reserved”(RAN)的方案:

1) 保留阶段

(a) 当产生的子个体的适应度大于父个体的适应度时,用子个体替代父个体;

(b) 否则以概率 $\exp(-\Delta f/E_p)$ 来替换父个体。 Δf 为父个体与子个体的适应度之差, E_p 作为染色体适应度值的均方差,由以下公式来计算: $E_p = \frac{1}{M} \sum_{i=1}^M (f(i) - \bar{f})^2$ (其中 f 为适应度函数)。在遗传算法执行的初期,由于个体间差异较大, E_p 较大,因而概率 $\exp(-\Delta f/E_p)$ 较大,此时的群体更新考虑更多的是保证染色体的多样性;在算法后期, E_p 较小,因而概率 $\exp(-\Delta f/E_p)$ 较小,保证优秀的父体被保留下来。

2) 不保留阶段

从原种群 Oldpop 中删除适应度最优的 40% 个体,然后重新构造新的路径染色体填补原种群的空缺,得到新种群 Newpop。同时将变异概率设为 1,交叉概率设为 0 来加强种群 Newpop 的全局勘探能力。

以上的设计是考虑到在传统的 GA 算法进化后期,染色体基本都收敛到某个最优解。在此基础上要想进一步改进是非常困难的,所以此时需要对种群进行一次“振荡”,以此来拓展解空间,从而跳出可能陷入的局部最优解空间,使之在最优解方向上做搜索。其中种群进行“振荡”的次数由振荡参数(oscillator)来控制。

总的说来,在进行种群更替时,算法首先进入保留阶段。而在该阶段的每一代演化结束后,都对当前群体中 40% 的最优个体的均值和上一代群体中 40% 的最优个体的均值进行比较。若适应度均值在进化了连续指定的代数后都没有增加,则强制转入不保留阶段。直到新种群 Newpop 再次收敛到一个适应度值,然后合并种群 Oldpop 和 Newpop,依概率 $\exp(-\Delta f/E_p)$ 决定保留哪些新个体构成新种群,从而替代 Oldpop 完成种群的进化(其中 N 为种群中的染色体个数)。

对于 oscillator 的取值问题,我们做了相关的实验。在实验中发现,RAN 策略随着 oscillator 的增大,其求解成功率有了很大的提高。特别是当 oscillator = 5 时,RAN 策略基本演变成采用双种群的 SA 策略^[16]。两者求得最优解概率几近相等,同时稳定性也较好。由此得出结论:通过增加 oscillator 的值,可以提高算法收敛到最优解的概率。在这里应当指出,我们没有就 oscillator 的上限值(即当 oscillator 大于某阈值后,求得最优解的概率大于一确定值)这个方向上研究下去,因为我们关心的是该策略的可用性,而不是在个别问题上取得的最好执行结果。同时考虑到 oscillator 值增长给算法平均收敛时间带来的计算复杂度,为了在收敛时间和求解成功率上获得一个有效的平衡,后面的实验取 oscillator = 3。

1.5 适应度函数

本文定义的 Steiner 树问题适应度函数为:

$$F(T(s, M)) = \sum \{ u_1 C(e) + u_2 \Psi(\Delta - D(e)) + u_3 \Psi(\delta - W(T)) \} + e \in T$$

其中网络延迟和其抖动最大值分别为 Δ 和 δ ; $u_i > 0$ 为加权因子 ($i = 1, 2, 3$); 单边费用函数 $C(e)$; 网络时延函数 $D(e)$; 时延抖动函数 $W(T)$; $\Psi(z)$ 为处罚函数, 定义如下:

$$\Psi(z) = \begin{cases} \varepsilon & z \leq 0 \\ 1 & z > 0 \end{cases}$$

其中参数 ε 作为惩罚度取 $[0, 0.5]$, 显然 $\min(F(T))$ 即所求解。

1.6 基因库的更新

每一代演化完后, 如果解的质量有所提高, 都有必要从现有群体中提取有效信息更新基因库, 从而使得基因库更好地指导后继的演化过程。更新的机理是: 1) 若染色体中产生了一条新的路径, 则将其加入到基因库对应终点的十字链表中; 2) 改变基因库中路径的优先被选择权, 将那些存在于当前最优解或次优解中的路径的权值提高。基因库被应用在 GP-GA 算法的 Disjoin 算子的子树缝合、Mutation 1 和下面介绍的动态 GP-GA 算法的路径选择中。

2 动态网络的 GP-GA 算法 GA-Online

在实际应用中, 组播的成员常随时加入或退出。如果重新建立组播树, 那么在时间和费用的消耗上无疑是巨大的。这时可以利用先前求得的最优染色体和基因库中保存的较优路径来构造初始当前种群, 加快算法的收敛速度。以下给出动态目标节点 t 的增加 JOIN(t) 和删除 REMOVE(t) 算法:

Procedure JOIN (t):

Begin

将节点 t 加入基因库

随机构造初始节点 s 到 t 的路径, 并调整染色体结构;

根据已经优化的染色体来构建初始种群 P ;

GP-GA ($N + t$);

End;

表 1 算法 A-GA^[7]、B-GA^[8] 和 GP-GA 求解效率分析

网络拓扑结构	算法平均求解成功率								
	100 代			300 代			500 代		
	A-GA	B-GA	GP-GA	A-GA	B-GA	GP-GA	A-GA	B-GA	GP-GA
$ N = 50, E = 150, (\varepsilon = 20\%)$	43.9%	33.6%	32.7%	76.8%	75.3%	75.8%	88.6%	90.3%	90.6%
$ N = 100, E = 300, (\varepsilon = 15\%)$	26.37%	26.62%	25.32%	54.3%	53.6%	70.8%	78.2%	74.88%	87.6%
$ N = 150, E = 500, (\varepsilon = 18\%)$	18.6%	14.3%	15.2%	48.65%	45.3%	66.7%	64.5%	68.7%	84.5%
$ N = 200, E = 725, (\varepsilon = 13\%)$	/	/	/	45.8%	47.2%	60.8%	50.1%	49.3%	82.1%

由表 1 可以看出, GP-GA 算法由于采用了保留与不保留的 (RAN) 进化策略, 在进化速度上要比 A^[7]、B^[8] 稍慢。但是由于基因库和特殊交叉算子, 收敛速度也处在一个可接受的范围内。当然 GP-GA 最大的优点就在于该算法平均求解问题的成功率始终保持较高的水平 (80% 以上), 而 A、B 随着网络规模的增加下降很快。可见 GP-GA 算法对于大中型拓扑的网络是十分适合的。

3.2 动态网络条件下的 GP-GA 算法

对于算法 A, T_A 表征算法 A 生成的组播树, $Cost(T_A)$ 是该生成树中所有链路费用和, 则定义算法 A 相对于算法 GA-Online 的“相对费用比”为: $\bar{E}_A = [Cost(T_A) - Cost(T_{GA-Online})] / Cost(T_{GA-Online})$ 。算法 MPH 定义见文献 [7]。实验结果如表 2。

从表 2 可看出, 动态 GA-Online 求得的最小 Steiner 树在费用上和 GA-Off 没什么区别, 两者都能很好地收敛到一个最优解 (而 MPH 算法相对而言在解稳定性和代价上都不是很好)。由此我们认为 GA-Online 能有效地屏蔽网络状态的非精确性, 较好地适应动态变化的网络环境。

M = BestOf (P);

Output M;

Procedure REMOVE (t):

Begin

从个体中删除与节点 T 有关的路径;

在先前得到的最优个体的基础上初始化种群 P ;

GP-GA ($N - t$);

End;

M = BestOf (P);

Output M;

3 仿真与分析

网络模型采用文献 [16] 提出的随机图模型, 该模型产生的随机图与真实网络比较接近, 同时文献 [7, 8] 也采用了同样的模型进行模拟实验。这样有利于消除环境因素给算法带来的实验误差, 客观地进行算法性能的比较。实验主要从以下两个方面来验证提出的 GP-GA 算法在解决 Steiner 问题时是否具有较好的健壮性和可行性。

3.1 算法平均求解成功率的比较

首先对几个参数进行解释: Pc 为交叉率, 取值 $0.2 \sim 0.9$; Pm 为变异率, 取值 $0.01 \sim 0.1$; $Popsiz$ 为种群个数, 取值 $[20, 100]$; $Maxgens$ 为最大世代数, 取值 $[2000, 4000]$; $oscillator = 3$ 。遗传算法的终止条件是: 所有种群都收敛到同一个解, 或算法迭代到了指定的最大代数。网络拓扑中 $|N|$ 为节点总个数, $|E|$ 为边个数, ε 为组播密度 (%), 延时和费用函数分别在 $[0, 50]$ ms 和 $[0, 200]$ 上正态分布。时延最大值 Δ 取 25, 时延抖动最大值 δ 取 6; 每组拓扑进行 100 次实验, 最后得到的结果如表 1 所示。

表 2 GP-GA 算法求解动态组播路由问题

网络拓扑结构	$\bar{E}_{MPH} (\%)$	$\bar{E}_{GA-Off} (\%)$	$\bar{E}_{GA-Online} (\%)$
$ N = 50, E = 150, (\varepsilon = 20\%)$	0	0	0
$ N = 50, E = 170, (\varepsilon = 20\%)$	3.08	0.01	0
$ N = 50, E = 185, (\varepsilon = 20\%)$	1.46	0	0
$ N = 100, E = 300, (\varepsilon = 15\%)$	4.06	0	0
$ N = 100, E = 350, (\varepsilon = 15\%)$	2.29	0.03	0
$ N = 100, E = 400, (\varepsilon = 15\%)$	0	0	0
$ N = 100, E = 470, (\varepsilon = 15\%)$	1.69	-0.12	0
$ N = 100, E = 500, (\varepsilon = 15\%)$	3.32	0.13	0
$ N = 150, E = 575, (\varepsilon = 18\%)$	4.10	-0.10	0
$ N = 150, E = 700, (\varepsilon = 18\%)$	0.72	0	0
$ N = 150, E = 780, (\varepsilon = 18\%)$	1.84	0	0
$ N = 150, E = 835, (\varepsilon = 18\%)$	2.55	0	0
$ N = 150, E = 900, (\varepsilon = 18\%)$	3.62	0	0
$ N = 200, E = 975, (\varepsilon = 13\%)$	0.06	0.14	0
$ N = 200, E = 1250, (\varepsilon = 13\%)$	1.16	0	0

(下转第 1952 页)

是当 $\lambda = 0.8$ 时的聚类结果优于 LF 算法的聚类结果。由实验结果还可看出,该算法还可进一步改进。蚁群算法的聚类数目有可能大于实际的聚类数目,如某些类之间很相近,可采用本算法或 FCM^[5] 等算法进行二次聚类。

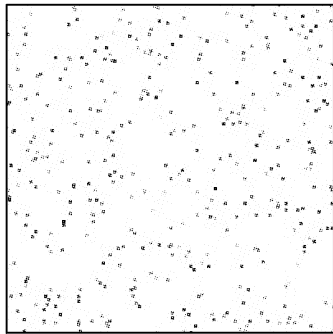


图3 400个彩色点随机分布于二维网格上

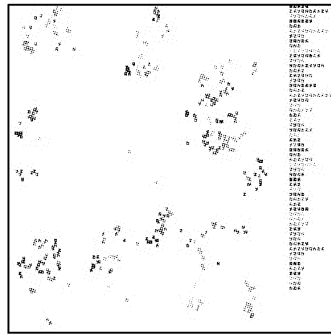


图4 $\lambda = 0.6$ 时的聚类结果

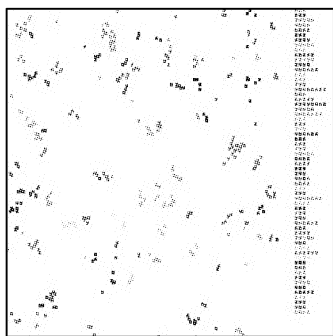


图5 $\lambda = 0.8$ 时的聚类结果

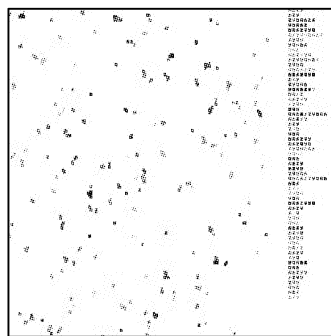


图6 LF 算法的聚类结果

4 结语

针对 LF 算法中存在的问题,改进后的算法更接近人们的思维习惯,并具有更好的稳定性。蚁群算法的研究时间较短,目前还没有形成较系统的分析方法和坚实的数学基础,各种参数的确定也没有一定的理论指导,但这也是下一步研究的重点。

参考文献:

- [1] 杨新斌,孙京浩,黄道. 一种进化聚类学习新方法[J]. 计算机工程与应用, 2003; 39(15): 60-62.
- [2] LUMER E, FAIETA B. Diversity and adaptation in populations of clustering ants[A]. Proceedings of third international conference on simulation of adaptive behavior: from animals to animates 3[C]. Cambridge, MA: MIT Press, 1994. 499-508.
- [3] WU B, SHI Z. A clustering algorithm based on swarm intelligence[A]. Proceedings IEEE international conferences on info-tech & info-net proceeding [C]. Beijing, 2001. 58-66.
- [4] 杨燕,靳蕃, KAMEL M. 一种基于蚁群算法的聚类组合方法[J]. 铁道学报, 2004; 26(4): 64-69.
- [5] KANADE PM, HALL LO. Fuzzy Ants as a Clustering Concept[A]. Proceedings of the 22nd International Conference of the North American Fuzzy Information Processing Society[C]. USA, 2003. 227-232.

(上接第 1949 页)

4 结语

本文提出基于基因库的遗传算法模型来解决网络中最小 Steiner 树问题,并针对传统方法求解过程中遇到的求解成功率不高的问题,提出了将传统 GA 和保留与不保留的(RAN)进化策略相嵌套的混合优化算法 GP-GA。同时考虑到进化策略可能带来的收敛速度慢的问题,我们又改进了交叉和变异算子。大量的仿真实验结果表明:采用此算法不论是在已知拓扑还是动态网络的条件下均可获得较高的求解概率,具有较大的实用价值。

参考文献:

- [1] 孙文生,刘泽民. 组播路由调度的神经网络方法[J]. 通信学报, 1998, 19(11).
- [2] 宋学军,刘巍. 多点并行蚁群搜索在多限制动态组播中的应用研究[J]. 电路与系统学报, 2004, 9(1).
- [3] 刘芳,冯小军. 免疫组播路由选择算法[J]. 计算机学报, 2003, 26(6).
- [4] 张琨,王珩,刘凤玉,等. 一种基于禁忌搜索的时延约束组播路由算法[J]. 计算机工程, 2005, 31(11).
- [5] Wang XW, Cheng H, Cao JN, et al. A Simulated-annealing based QoS multicasting algorithm[A]. Proceedings of International Conference on Communication Technology[C]. Beijing: IEEE, 2003. 469-473.
- [6] ESBENSEN H. Computing Near-Optimal Solutions to the Steiner Problem in a Graph Using a Genetic Algorithm[J]. Networks, 1995, 26: 173-185.
- [7] CHEN L, YANG ZY, XU ZQ. A Degree-Delay-Constrained Genetic Algorithm for Multicast Routing Tree[A]. Proceedings of the Fourth

International Conference on Computer and Information Technology (CIT'04)[C]. IEEE, 2004.

- [8] HAGHIGHAT AT, FAEZ K, DEHGHAN M, et al. A Genetic Algorithm for Steiner Tree Optimization with Multiple Constraints Using Prüfer Number[A]. Proceedings of the First EurAsian Conference on Information and Communication Technology[C]. Lecture Notes In Computer Science, 2002, Vol 2510: 272-280.
- [9] 陈明,李志杰. 基于遗传算法的实时组播通信路由算法[J]. 软件学报, 2001, 12(5).
- [10] 石坚,邹玲,董天临,等. 遗传算法在组播路由选择中的应用[J]. 电子学报, 2000, 5(28).
- [11] 陈佳,康立山,陈毓屏. 用遗传算法解带时延及时延抖动约束的组播路由优化问题[J]. 计算机工程与应用, 2003, (21): 147-149.
- [12] KARABI M, FATHY M, DEHGHAN M. QoS Multicast Routing Based On a Heuristic Genetic Algorithm[A]. 17th Canadian Conference on Electrical and Computer Engineering[C]. Niagara Falls, Canada, 2004. 1727-1730.
- [13] ESBENSEN H. A Genetic Algorithm for the Steiner Problem in a Graph[J]. IEEE Journal on Selected Areas in Communications, 1994: 402-406.
- [14] ZHOU XW, CHEN CJ. A Genetic Algorithm for Multicasting Routing Problem[A]. International Conference Communication Technology Proceedings, WCC-ICCT[C]. Beijing: IEEE, 2000. 1248-1253.
- [15] 陈品,刘三阳. 基于双种群遗传策略的组播路由算法[J]. 电子与信息学报, 2002, 12(24).
- [16] SALAMA HF, REEVES DS, VINIOTIS Y. Evaluation of multicast routing algorithms for real-time communication on high-speed networks, IEEE Journal on Selected Areas in Communications, 1997, 15(3): 332-345.