

DNA 遗传算法的 QoS 多播路由优化

唐天兵¹, 申文杰¹, 韦凌云²

(1. 广西大学计算机与电子信息学院, 南宁 530004; 2. 北京邮电大学自动化学院, 北京 100876)

摘要: 针对网络中的 QoS 多播路由问题, 提出一种基于 DNA 的混合遗传算法, 以遗传算法为框架, 采用 DNA 编码方式, 对交叉算子和变异算子进行改进, 同时把小生境技术和模拟退火融入框架中, 使小生境中的个体独立进行模拟退火操作, 达到维护种群多样性和增强局部搜索能力的目的。实验结果验证了该算法的有效性。

关键词: 多播路由; 遗传算法; DNA 计算; 小生境; 模拟退火

QoS Multicast Routing Optimization for DNA Genetic Algorithm

TANG Tian-bing¹, SHEN Wen-jie¹, WEI Ling-yun²

(1. School of Computer, Electronics and Information, Guangxi University, Nanning 530004;

2. School of Automation, Beijing University of Posts and Telecommunications, Beijing 100876)

【Abstract】 A hybrid Genetic Algorithm(GA) based on DNA is proposed for the optimization of multicast routing with Quality of Service(QoS) constraint. Based on the framework of GA with DNA coding and improved crossover and mutation operators, the proposed algorithm is also combined with improved simulated annealing and niche operation to make the individuals of a niche enforce simulated annealing operation independently and to keep the diversity of population, and further to improve search capability of the algorithm and convergence rate. Experimental results demonstrate this algorithm is viable and efficient.

【Key words】 multicast routing; Genetic Algorithm(GA); DNA computing; niche; simulated annealing

1 概述

随着因特网规模的不断扩大, 网上的实时业务也随之增长, 实时业务对网络传输的代价、带宽、丢包率、延时和延时抖动等特性较敏感, 因此, 在网络上采用 QoS 服务确保实时业务的通信质量。QoS 多播路由优化是实现 QoS 服务的有力保障。QoS 多播路由优化可以看成是从源节点出发, 遍历所有目的节点, 找到一棵满足所有约束条件且代价最小的 Steiner 树问题, 所以, QoS 多播路由优化是一个 NP 完全问题^[1]。遗传算法是解决 NP 完全问题的有效方法。遗传算法(Genetic Algorithm, GA)是种以 Darwin 自然进化论与 Mendel 遗传变异理论为基础的求解复杂全局优化问题的仿生型算法, 已在众多领域获得应用, 但它存在的显著不足是容易过早收敛和局部搜索能力差^[2-3], 解决办法之一是把 DNA 计算引入到遗传算法中。文献[4]提出 DNA 计算的概念, DNA 携带丰富的遗传信息, 是重要的遗传物质。DNA 计算是利用 DNA 分子的各种生化反应来完成计算, 而遗传算法是一种在分子水平模拟生物进化过程来求解复杂问题的有效算法, 两者具有某种必然的联系。解决办法之二是在遗传算法中引入模拟退火等较强的局部搜索算法^[5]。

针对 QoS 多播路由优化, 本文构造一个基于 DNA 的混合遗传算法(DNA-Hybrid Genetic Algorithm, DNA-HGA)。该算法在微观遗传策略方面, 采用 DNA 编码, 改进交叉操作和变异操作; 在宏观遗传策略方面, 执行遗传算法的基本操作之后, 再对小生境淘汰操作后的保留个体进行改进的模拟退火操作。通过和其他文献进行比较, 本文算法有效、收敛性好。

2 QoS多播路由的优化模型

QoS 多播路由优化问题是在网络中寻找一个满足多 QoS 约束的最优多播树。网络可以表示成一个带权的无向图 $G=(V, E)$, 其中, V 表示网络中的节点(路由器); E 表示边集(通信链路); $|V|$ 是节点的个数; $|E|$ 是边的个数。假设 s 是源节点, M 是目的节点的集合, T 是源节点 s 到所有目的节点 $d_i \in M$ 的所有多播树, t 是 T 中任意一个多播树, P 是树 T 中 s 到某个目的节点的路径, $e(v_i, v_j) \in t$ 是多播树的一条边, $n \in V$ 是任意的一个节点。式(1)~式(4)分别表示多播树中一条路径的带宽、延时、延时抖动和丢包率。

$$\text{bandwidth}(P) = \min\{\text{bandwidth}(e), e \in P\} \quad (1)$$

$$\text{delay}(P) = \sum_{e \in P} \text{delay}(e) \quad (2)$$

$$\text{delay_jitter}(P) = \sum_{e \in P} \text{delay_jitter}(e) \quad (3)$$

$$\text{loss}(P) = 1 - \prod_{n \in P} (1 - \text{loss}(n)) \quad (4)$$

假设最优多播树的最小带宽约束为 B 、最大延时为 D 、最大延时抖动为 DJ 和最大丢包率为 Q , QoS 多播路由优化以

基金项目: 国家自然科学基金资助项目(50605010); 广西教育厅科研基金资助项目(200911LX15); 广西研究生教育创新计划基金资助项目(105930903042)

作者简介: 唐天兵(1972—), 男, 副教授、硕士, 主研方向: 并行分布式计算, 优化算法; 申文杰, 硕士研究生; 韦凌云, 副教授、博士后

收稿日期: 2009-11-20 **E-mail:** tbtang@gxu.edu.cn

多播树的边为设计变量，以代价最小为目标，以带宽约束、延时约束、延时抖动约束和丢包率约束为约束条件，优化模型如下：

$$\min_{e \in E} \text{cost}(t(s, M)) = \sum \text{cost}(e) \quad (5)$$

$$\text{s.t. } \min_{d_i \in M} \{\text{bandwidth}(P(s, d_i))\} \geq B \quad (6)$$

$$\max_{d_i \in M} \{\text{delay}(P(s, d_i))\} \leq D \quad (7)$$

$$\max_{d_i \in M} \{\text{delay_jitter}(P(s, d_i))\} \leq DJ \quad (8)$$

$$\max_{d_i \in M} \{\text{loss}(P(s, d_i))\} \leq L \quad (9)$$

其中，式(5)是多播树的代价；式(6)是带宽约束；式(7)是延时约束；式(8)是延时抖动约束；式(9)是丢包率约束。

3 算法设计

3.1 编码

编码是遗传算法要解决的首要问题。之前多播遗传算法的编码方法主要有： $N \times N$ 二进制矩阵表示法^[6]，路径表示法^[7]，非形式化树表示法^[8]等。本文参考文献[4]的编码方法，引入 DNA 编码方式，主要考虑网络节点和链路如何编码，一是节点编码由 {A, T, C, G} 4 个碱基组成，每个节点采用长度为 20 个碱基寡聚核苷酸片段表示，其中，节点编码的前 2 个和后 2 个碱基限制为 {T, C, G} 上的字母，便于区分节点和链路编码。记为 $N_i (i=1, 2, \dots, n)$ ；二是链路编码由前一个节点 N_i 的后 10 个寡聚核苷酸片段和后一个节点 N_j 的前 10 个寡聚核苷酸片段组成，为了便于表示一条路径，源节点和目的节点采用自身 20 个寡聚核苷酸片段表示，这样可以保证节点 N_i 与 N_j 的相互可达性和路径的方向性。种群初始化方法：首先选择一条从源节点到任意一个目的节点的路径加入到多播树中，路径中的节点记为 node，对于剩余的目的节点，分别选择一条从目的节点到 node 中的一个节点的一条路径，加入到多播树中，同时将各路径中新增节点加入到 node 里面，这样可以保证构建一棵不含环路的最优多播树，算法如下：

假设源节点为 s ，目的节点集合为 E

输出 多播树 T

for $i=1:\text{length}(E)$

if $i==1$

 任意选择一个目的节点 n ;

 temp= s ;

 node= s ;

 while temp $\neq n$

 任意选择一条与 s 连接的链路 e ，连接的节点为 p ;

 path=path $\cup e$;

 node=node $\cup p$;

 temp= p ;

 end

 T=path;

else

 从剩余的目的节点中选择一个节点 n ;

 temp= n ;

 while temp \notin node

 任意选择一条与 temp 连接的链路 e ，连接的节点为 p ;

 path=path $\cup e$;

 if $p \notin$ node

 node=node $\cup p$;

 end

 temp= p ;

 end

T=path;

end

end

3.2 适应度计算

适应度函数直接影响遗传算法的性能。适应度较高的个体遗传到下一代的概率就较大，而适应度较低的个体遗传到下一代的概率就较小。文中构造的适应度函数表达式如下：

$$F = \frac{1}{C_i + M \sum_{i=1}^n \text{cons}_i + N} \quad (10)$$

其中， C_i 为目标函数多播树的代价； cons_i 为第 i 个约束违反量的绝对值； M 为惩罚因子，设为很大的数（如 $M = 1\ 000\ 000$ ）， N 为一个很小的数（如 $N = 0.000\ 01$ ），以防止被 0 除。

3.3 交叉算子和变异算子

交叉操作是遗传算法的主要操作，在交叉过程中，首先从种群中任选 2 个个体，根据这 2 个个体的所有路径重新形成一个网络图，再根据种群初始化方法重新生成 2 个个体。此方法可以保证子代都是来自父代基因，不会形成非法的多播树，但可能会使生成的个体跟交叉前的个体一样。为了改善这种情况，当个体没有发生变化时，选择一个个体进行变异操作。

变异操作是为了保持种群的多样性。对于进行变异的个体，可以从多播树中的一个目的节点出发，删除一条路径，之后再从该目的节点重新寻找一条路径添加到树中。

文中采用自适应交叉算子和自适应变异算子，交叉概率 p_c 和变异概率 p_m 能够随着个体的适应度自动变化。其思想是当种群适应度比较集中或局部最优时，使 p_c 和 p_m 增大，当种群适应度比较分散时，使 p_c 和 p_m 减少。因此， p_c 和 p_m 按如下公式自适应调整：

$$p_c = \begin{cases} k_1 \frac{f_{\max} - f'}{f_{\max} - f_{\text{ave}}} & f' \geq f_{\text{ave}} \\ k_2 & f' < f_{\text{ave}} \end{cases} \quad (11)$$

$$p_m = \begin{cases} k_3 \frac{f_{\max} - f}{f_{\max} - f_{\text{ave}}} & f \geq f_{\text{ave}} \\ k_4 & f < f_{\text{ave}} \end{cases} \quad (12)$$

其中， f_{\max} 表示种群的最大适应度； f_{ave} 表示种群的平均适应度； f' 是交叉的 2 个个体中较大的适应度； f 是变异个体的适应度。

3.4 模拟退火操作

模拟退火是一种随机搜索技术，其基本思想是，从一个初始当前解开始，在当前解的邻域内，以一定的概率 P 寻找一个非局部最优解，并令此解作为当前解继续查找，从而避免陷入局部最优。将模拟退火操作应用到遗传算法的局部搜索过程中，关键点是如何在当前解的邻域内，随机地产生一个新解，文中采用类似变异操作的扰动策略产生新解。

3.5 算法流程

根据上述思想，基于 DNA 混合遗传算法的 QoS 多播路由优化算法实现如下：

(1) 初始化种群。设置进化代数计数器 $t=1$ ，最大进化代数 T 。采用 DNA 编码，通过上述初始化方法生成由 M 个个体组成的初始种群 $P(t)$ ，计算各个个体的适应度，将 $P(t)$ 作为父代种群。根据个体的适应度进行降序排序，记忆前 N 个

个体($N < M$)。

(2)选择运算。采用随机联赛选择方式,对种群 $P(t)$ 进行选择运算,得到种群 $P(t)'$ 。

(3)自适应交叉操作。采用自适应的 p_c 对种群 $P(t)$ 进行的交叉操作,得到种群 $P(t)''$ 。

(4)自适应变异操作。采用自适应的 p_m 对种群 $P(t)''$ 进行变异操作,得到种群 $P(t)'''$ 。

(5)小生境生成操作。将第(4)步得到的 M 个个体和所记忆的 N 个个体合并成 $M+N$ 个个体的新群体,对这 $M+N$ 个个体,按照式(13)和式(14)计算每 2 个个体 X_i 和 X_j 之间的海明距离,小生境半径为 D ,对每一个个体 X_i ,当其适应度 $F(X_i) > 0$ 时,若 $\|X_i - X_j\| < D$ 时,则惩罚个体 X_i ,令 $F(X_i) = 0$,得到种群 $P(t)''''$ 。

$$f(x, y) = \begin{cases} 1 & \text{若 } x = y \\ 0 & \text{若 } x \neq y \end{cases} \quad (13)$$

$$\|X_i - X_j\| = \sum_{k=1}^n f(x_{ik}, y_{jk}) \quad (14)$$

(6)模拟退火操作。对种群 $P(t)''''$ 中适应度不为 0 的个体进行改进的模拟退火操作,降序排序之后取前 M 个个体形成种群 $P(t)$ 作为下一代进化种群。

(7)终止条件判断。若 $t \leq T$,更新进化迭代计数器 $t=t+1$,转到第(2)步,若 $t > T$,则输出结果,计算结束。

4 算例

本文采用 Matlab7.0 进行仿真实验。遗传操作的参数设置为:种群规模为 30,遗传迭代次数为 200,其中,式(2)和式(3)中, $k_1=0.85$, $k_2=1$, $k_3=0.5$, $k_4=1$ 。模拟退火操作的参数设置为:初始温度 $t_0=30^\circ\text{C}$,终止温度 $t_f=0.05^\circ\text{C}$,模拟退火的迭代次数 $step=10$,温度衰减控制参数 $L=0.5$ 。算法分别与文献[1]和文献[5]进行比较。文献[1]的多播 QoS 约束取链路时延 Δ 和带宽 β 这 2 个参数,分别取 $\Delta=9$, $\beta=8$ 和 $\Delta=10$, $\beta=9$ 。图 1 和图 2 分别是 $\Delta=9$, $\beta=8$ 和 $\Delta=10$, $\beta=9$ 与文献[1]的比较情况。可以看出,本文的代价最优值与文献[1]中的计算结果完全一致,但本文在 20 代左右就接近代价最优值,明显少于文献[1]算法,说明本文算法效率高、收敛性好。

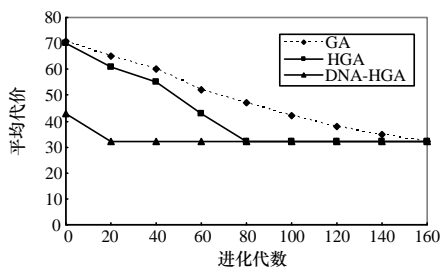


图 1 当 $\Delta=9$, $\beta=8$ 时的收敛情况

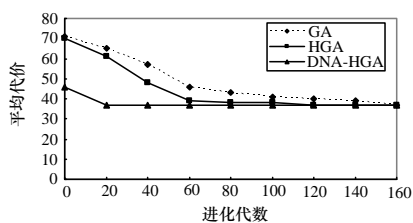


图 2 当 $\Delta=10$, $\beta=9$ 时的收敛情况

与文献[5]进行比较时,本文算例主要考虑带宽、延时和延时抖动 3 个约束。本文采用改进的 Salama 算法随机生成网络拓扑图,设链路代价是 2~10 之间的随机数,带宽是 50 Kb/s~200 Kb/s 之间的随机数,链路延时等于链路距离除以 2/3 光速,延时抖动是 5 ms~15 ms 之间的随机数,丢包率是 0~0.01 之间的随机数。

图 3 列出了在不同网络节点与多播树的最小代价之间的比较。可以看出,本文算法明显优于文献[5]的结果,进一步验证了该算法的有效性。

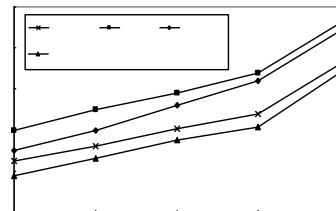


图 3 代价性能比较

5 结束语

本文以遗传算法为主体,采用 DNA 编码方式,有机融入小生境技术和模拟退火操作,达到维护种群多样性和增强局部搜索能力的目的。通过和其他文献进行比较可知,本文算法可以节约路由搜索时间和减少路由开销,为解决多约束 QoS 多播路由问题提供一个新方法。

下一步工作是研究该混合遗传算法的并行化和移植到动态多播路由问题。

参考文献

- [1] 范一鸣,余建军,方智敏.融合小生境机制的 QoS 多播路由遗传模拟退火算法[J]. 通讯学报, 2008, 29(5): 65-71.
- [2] 徐宗本,高勇.遗传算法过早收敛现象的特征分析及其预防[J]. 中国科学: E 辑, 1996, 26(4): 364-375.
- [3] 黎明,龙佳丽,盛伟翔.遗传算法群体规模的研究[J]. 计算机工程, 2009, 35(1): 162-164.
- [4] Adleman L M. Molecular Computation of Solutions to Combinatorial Problems[J]. Science, 1994, 266(5187): 1021-1024.
- [5] Zhang Li, Cai Lianbo, Li Meng, et al. A Method for Least-cost QoS Multicast Routing Based on Genetic Simulated Annealing Algorithm[J]. Computer Communications, 2009, 32(1): 105-110.
- [6] Luo Junzhou. QoS Routing Based on Genetic Algorithm[J]. Computer Communications, 1999, 22(9): 1394-1399.
- [7] Sun Baolin, Pi Shangchao, Gui Chao, et al. Multiple Constraints QoS Multicast Routing Optimization Algorithm in Manet Based on GA[J]. Progress in Natural Science, 2008, 18(3): 331-336.
- [8] Liu Jianhua, Yan Dekun. Niche Genetic Algorithm with Accurate Optimization Performance[J]. Journal of China University of Mining and Technology, 2005, 15(2): 100-104.

编辑 陈文