

基于遗传算法的道路网综合模型

邓红艳^{1,2,3} 武芳¹ 翟仁健¹ 刘薇薇¹

(1 信息工程大学测绘学院, 郑州市陇海中路66号, 450052)

(2 辽宁工程技术大学地理信息系统实验室, 阜新市中华路47号, 123000)

(3 空军指挥学院, 北京市海淀区, 100089)

摘要: 提出了一种基于遗传算法的道路网综合模型。实验表明, 该模型在保持道路网的空间分布性和实用性方面具有良好的效果。

关键词: 道路网综合; 遗传算法; 拓扑; 适应度函数

中图法分类号: P283.1

道路网作为地图上大量存在的要素一直是制图领域研究的重点, 自动制图综合的研究也不例外。对道路网的综合, 已经有许多学者^[1-3]进行了研究, 虽然没有从根本上解决该问题, 但也取得了一定的成绩。早期的道路网综合模型主要是基于道路等级的简单选取。现代的道路网综合模型主要是指基于图论的道路选取, 这种方法虽然在建立拓扑关系、确定道路节点权值方面具有一定困难, 但是依据图论的一些基本特性, 它能在一定程度上保持道路网的分布特性。

如何既能使综合后的道路网络保持原有道路的空间分布特性, 又具有良好的经济和实用价值, 无论对于普通地图制图综合还是支持地理信息系统的数据库综合都具有重要意义。其实质就是一个多目标的优化问题, 采用一般的优化算法是很难实现的, 因此, 需要结合道路网综合的基本原则寻求一种新的方法(本文将以最短路程作为经济实用价值的考虑因素)。

本文基于遗传算法的基本原理, 从全局的角度(考虑道路的空间分布特性与经济实用价值) 提出了一种基于遗传算法的道路网综合新方法。

1 道路网综合的基本原则

道路网的综合过程主要包括道路的自动选取、道路形状的自动化简和道路的自动概

括^[2,4,5]。现阶段道路网的综合主要是指道路网的选取。为保持制图区域原有的地理特点和满足地图的用途要求, 道路网综合应考虑以下因素。

1) 保持道路网及其相关要素的拓扑特征。道路与居民地之间存在着有机的联系, 即道路的取舍要考虑与居民地的贯通性和一致性。

2) 保持道路网图形几何特征。

3) 保持道路网分布特征。道路网综合后, 应保持原有的道路分布特性。同时, 选取后, 由于道路与道路之间的距离变小, 符号化时可能会出现冲突, 即符号压盖。这在实际操作中主要通过线条要素之间的关系处理完成。

2 基于遗传算法的道路网综合模型

道路网综合时, 要最大限度地保持道路网的网眼分布、排列规律、经济实用价值等因素。基于遗传算法的道路网综合模型的基本过程是: 首先, 对道路网数据与居民地数据进行预处理, 生成拓扑关系, 建立道路的几何分布特性, 同时计算道路网络中各居民地之间的最短路径, 作为与选取后比较的基准; 其次, 依据局部搜索策略和道路的几何分布特性生成多种可行的综合方案, 再根据适应度函数对每种方案进行经济实用价值的评估(本文经济实用价值主要考虑道路的通达性); 最后, 利用遗传算法的各种遗传算子^[6-8], 不断遗传

迭代, 获取在空间分布特性和经济实用价值上的良好结果。

2.1 编码的确定

编码是应用遗传算法要解决的首要问题, 它是在问题的候选范围内随机地产生一组候选解作为初始群体。

在道路选取时, 每条道路都有两种状态: 选取或者不选取。因此, 应选择编码、解码以及交叉运算都很方便的二进制编码方法^[6-8], 用 1 表示道路被选取, 0 表示不被选取, 编码长度为所有的道路数目。同时, 由于道路网是带有属性的特殊折线, 因此, 还应该考虑一些特殊情况, 例如等级高或者具有特殊意义的道路(如行军的惟一通道)都应该被选取。这些情况反映到染色体编码上为约束性条件: 等级高或者具有特殊意义的道路对应的染色体上等位基因始终为 1。

在编码时, 编码的长度等于道路网中的道路数目 N , 染色体等位基因的取值反映相应位置道路是否被选取: 1 表示被选中, 0 表示未被选中。其中一些等级非常高或者具有特殊意义的道路, 其编码取值始终为 1, 表示这些道路始终被选取。如对于道路数据 $\{L_1, L_2, L_3, L_4, L_5, L_6, L_7, L_8\}$, L_3, L_5 具有重要的意义, 这两位编码将始终为 1, 如果编码最后结果为 01111001, 则最后选择结果为 $\{L_2, L_3, L_4, L_5, L_8\}$ 。

2.2 适应度函数的确定

适应度是群体中个体对其环境的适应程度。适应度函数是为了求取个体的适应度而根据具体的应用目标构造的对适应度进行评价的函数, 它是遗传算法中一个很重要的控制参数, 类似于优化问题中的目标函数。一般情况下, 适应度函数与问题所要达到的要求密切相关, 并且要求尽量简单。针对道路网综合, 由于既要考虑道路的空间分布特性, 又要考虑道路的经济实用价值, 因此, 不但计算不便, 也存在着多目标函数求解困难的问题。本文将道路网络中所要达到的一些难以求取最优解, 但是可以求取最优近似解的主要目标设计适应度函数, 而将可以直接进行量化求解的过程作为局部优化策略代替遗传算子的变异算子。这样做可以有效地缩小遗传算法的搜索空间, 提高遗传算法的效率。衡量道路网络的经济实用价值与空间分布特性, 尽量将保持道路网的经济实用价值作为主要目标建立适应度函数:

$$F = \sum_{k=1}^{n_i} \sum_{i=1}^{n_j} d(P_k, P_j), \quad k \neq j \quad (1)$$

$$d(P_k, P_j) =$$

$$\begin{cases} 1, & D(P_k, P_j) \leq \varepsilon \quad \min D(P_k, P_j) \\ 0, & D(P_k, P_j) > \varepsilon \quad \min D(P_k, P_j) \end{cases} \quad (2)$$

其中, P_k, P_j 表示道路网络所覆盖的居民地抽象出来的节点; n_i 表示居民地节点的个数; $D(P_k, P_j)$ 表示道路选择后的两居民地节点之间的距离; $\min D(P_k, P_j)$ 表示道路选择前的两居民地节点之间的最短路径; ε 表示保持经济实用价值(两居民地之间的距离不太远)的程度, ε 是可变的。依据区域的不同, ε 取值不同: 当道路比较崎岖时, ε 取值较小, 表示该地区绕行困难; 当道路比较平坦时, ε 取值较大, 表示该地区行进轻松。 $d(P_k, P_j)$ 表示点 P_k, P_j 之间是否通行顺利的评价指数, $D(P_k, P_j) \leq \varepsilon \quad \min D(P_k, P_j)$ 表示综合后的道路网络两居民地之间的距离与综合前道路网之间的距离相差不远, 道路通行顺利, 取值为 1; $D(P_k, P_j) > \varepsilon \quad \min D(P_k, P_j)$ 表示综合后的道路网络两居民地之间的距离与综合前道路网之间的距离相差太远, 取值为 0, 表示选择后道路通行不顺利。

适应度函数 F 为所有 $d(P_k, P_j)$ 之和, F 越大, 表示需要绕行的道路越少, 适应度越高, 该道路网综合方案的经济实用价值越高; 反之则表示该道路网综合方案的经济实用价值越低。

2.3 局部搜索策略

由于本文采用的是遗传算法与局部搜索策略相结合的混合遗传算法, 因此, 在算法的实现过程中, 还涉及到局部搜索策略的确定。将道路网综合过程中保持空间分布特性这一目标作为次要因素进行考虑, 使得所有将要进行下一步操作的染色体^[6-8](备选方案)均是可行解, 即所有进行经济实用价值判断的道路网选择方案均保持空间分布特性。局部搜索策略如下。

1) 对一条有待进行局部搜索策略的染色体进行解码, 以确定原始道路网中各条道路的选择情况;

2) 从知识库中读取预先设置的道路等级;

3) 根据道路等级进行染色体的修正, 对等级很高和具有特殊意义的道路, 相应染色体上的取值始终设为 1, 如果为 0 则变成 1;

4) 在已经修正过的染色体基础上, 依据道路的选取情况, 构建新的拓扑关系;

5) 根据编图规范对道路网眼大小进行判断, 如果某道路网眼过小, 则构成该道路网眼的最低等级道路染色体上相应的等位基因取值变为 0; 如果某道路网眼太大, 看该道路网眼中是否存在没有被选取的道路, 如果存在, 将其中的最高等级

道路染色体上相应的等位基因变为 1, 直到所有的道路网眼均符合编图规范要求, 即道路网的空间分布状况被保持;

6) 该局部搜索策略不但用于染色体的初始产生过程, 保证最初参加遗传操作的染色体在保证几何空间分布这一目标上均是可行解, 也用于代替染色体的变异操作, 保证在遗传迭代过程中, 染色体在保证几何空间分布这一目标上也均是可行解。

实验证明, 采用遗传算法与局部搜索策略相结合, 利用启发式信息及相关领域的知识策略, 可避免在非可行解上浪费时间, 大大提高了系统的运行效率。

2.4 终止条件的设定

为了避免基于遗传算法的道路网综合模型在实际操作过程中出现不收敛的问题, 本文采用了两种终止条件: 确定遗传代数; 将平均适应度与最大适应度相结合进行判断, 如果平均适应度与最大适应度之差在限差范围内, 则运算终止, 表明遗传算法不再向好的方向优化。这两种方法结合进行, 无论达到哪种条件, 运算均终止。

综上所述, 基于遗传算法的道路网综合的总体步骤如下。

- 1) 根据给定的道路数据进行拓扑, 计算各居民地之间的最短路径, 作为与选取后比较的基准;
- 2) 根据给定的道路等级对道路进行分类;
- 3) 用局部搜索策略从高等级到低等级的顺序进行判断选择, 对于同一等级的道路按照随机顺序进行选择;
- 4) 进行遗传操作, 其中变异算子采用局部搜索策略代替;
- 5) 重新生成道路拓扑关系, 计算新的道路网络中各居民地之间的最短路径并与初始距离进行比较, 求出适应度函数;
- 6) 比较适应度, 生成新的群体。重复步骤 3) ~ 步骤 6), 直到满足终止条件;
- 7) 对最后一代群体中具有最高适应度的染色体进行解码处理, 以获取最终结果。

3 实验结果与结论

本文利用基于遗传算法的道路网综合模型, 对某一地区的道路网数据进行了综合实验。其中, 图 1 表示待综合的道路网络数据, 图上标记的居民地 A、B、C、D 表示需要考虑的居民地(为了说明清楚, 特选取了较少的居民地)。图 2 为综合后的道路网

数据, 从结果来看, 该方法的特点如下。

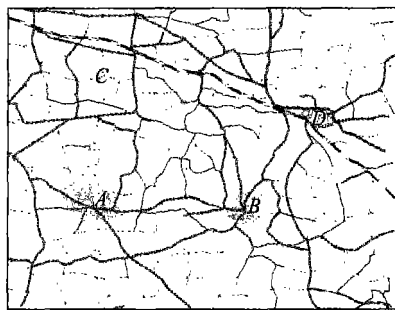


图 1 综合前的道路网数据

Fig. 1 Road Networks Data Before Generalization

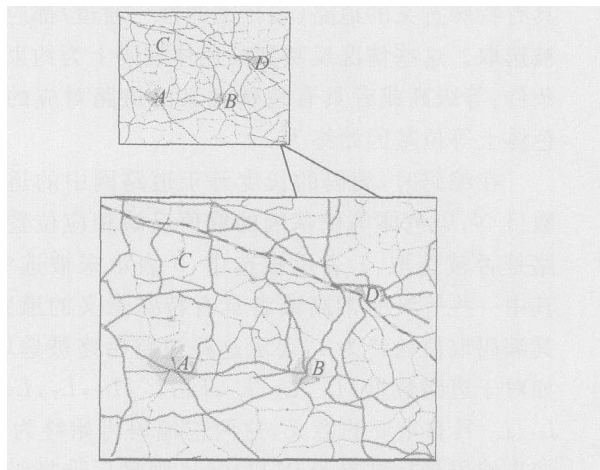


图 2 综合后的道路网数据

Fig. 2 Road Networks Data After Generalization

- 1) 保持密度分布特征及其分布规律, 即基本保持道路网的网眼特征;
- 2) 在考虑道路的空间分布特性的基础上, 最大程度地保证综合结果的经济适用性质, 即居民地之间的距离没有太大改变。

本文只是遗传算法用于道路网综合的一个初步尝试, 真正要用此模型来完全解决制图综合中的道路网问题还需要其他相关模型算法的完善, 如最短路径算法、拓扑关系生成算法等。

参 考 文 献

- [1] 武芳. 地图设计与编绘[M]. 郑州: 解放军测绘学院, 1997
- [2] 王家耀, 武芳. 数字地图自动制图综合原理与方法[M]. 北京: 解放军出版社, 1997
- [3] 祝国瑞, 郭礼珍, 尹贡白, 等. 地图设计与编绘[M]. 武汉: 武汉大学出版社, 2001
- [4] 王光霞. 自动制图综合系统的研究及居民地道路自动综合的实现[D]: [学位论文]. 郑州: 解放军测绘学院, 1994
- [5] 李敏强, 寇纪淞, 林丹, 等. 遗传算法的基本原理与

- 应用[M]. 北京: 科学出版社, 2002
- [6] 张文修, 梁怡. 遗传算法的数学基础[M]. 西安: 西安交通大学出版社, 1999
- [7] 蔡少华. GIS 图形空间关系的研究与实践[D]: [学位论文]. 郑州: 解放军测绘学院, 1999
- [8] Jiang B. A Structure Approach to the Model Generalization of an Urban Street Network[J]. *Geoinformatica*, 2004(2): 157-171
- [9] Duchêne C. Road Generalization Using Agents[C]. 9th Annual Conference on GIS Research in United Kingdom, Glamorgan, 2001
- [10] Yu Xuchu. Road Network Simplification with Knowledge-based Spatial Analysis[J]. *Supplement Journal of Geographical Sciences*, 2001(11): 54-62

第一作者简介: 邓红艳, 博士生, 主要研究方向是空间数据多尺度表达与 GIS。

E-mail: denghongyan_gis@163.com

A Generalization Model of Road Networks Based on Genetic Algorithm

DENG Hongyan^{1, 2, 3} WU Fang¹ ZHAI Renjian¹ LIU Weiw ei¹

(1 Institute of Surveying and Mapping, Information Engineering University, 66 Middle Longhai Road, Zhengzhou 450052, China)

(2 Laboratory of GIS, Liaoning Technical University, 47 Zhonghua Road, Fuxin 123000, China)

(3 Air Force Command College, Haidian District, Beijing 100089, China)

Abstract: A new generalization model of road networks based on genetic algorithm is proposed. Experimental results indicate that the principle of the method is correct and it performs well for keeping the practicability and spatial distribution characteristics of road networks.

Key words: generalization of road networks; genetic algorithm; topologic; fitness function

About the first author: DENG Hongyan, Ph. D candidate, majors in multi-scale representation and GIS.

E-mail: derghongyan_gis@163.com

本部相关期刊 2005 年引证情况公布

据《中国学术期刊综合引证 2005 年度报告》报道, 我部所属期刊引文计量情况如下:

	总被引频次	影响因子	即年指标	Web即年下载率
武汉大学学报·信息科学版	811	0.851	0.107	36.1
测绘信息与工程	162	0.326	0.038	23.3
地球空间信息科学学报	29	0.116	0	8.1

本报告是由中国科学文献计量评价研究中心完成的。