

一个新的无向图画图算法*

黄竞伟 康立山 陈毓屏

(武汉大学软件工程国家重点实验室 武汉 430072)

E-mail: jwhuang@public.wh.hb.cn

摘要 将一般无向图的画图问题转化为函数优化问题,用遗传算法求目标函数的最优解的近似值,从而得到无向图自动画图算法的一个一般框架.新方法的特点是:不同的画图算法的框架都一样,所不同的只是反映无向图画图问题的美观标准的目标函数.其优点在于,算法统一、方法简单、容易实现、便于修改,并且易于并行化,可以直接用来画非连通图.

关键词 无向图,画图,美观准则,算法,遗传算法.

中图法分类号 TP301

图是一种应用非常广泛的数据结构,它可以用来表示各种复杂的系统模型,其中顶点表示系统中的元素,边表示元素之间的关系.将图整齐、美观地画出来,对于理解和分析模型都十分重要.通常,表示复杂系统模型的图的顶点数和边数都相当大,以致于用人工方法将图整齐、美观地画出来几乎是不可能的,因而如何用计算机自动地将图整齐、美观地画出来便成为一个十分重要的课题.除此之外,画图(graph drawing)在电路布线、网络管理、软件工程、图形学、程序可视化、计算机辅助教学等领域中都有着众多的应用.近年来,画图受到了越来越多的关注,形成了计算机科学领域中的一个新的研究方向.

自1992年以来,国际上已举办了七届(每年一次)画图学术研讨会.文献[1]是一篇有关画图的综述性文章.通常,对各种不同类型的图有不同的画图算法,譬如有画树、平面图、有向图、无向图的算法.而在每一类图中,又根据图的特性和美观准则的不同,而有不同的算法.对于无向图画图算法,所使用的方法有弹性模型方法^[2]和力导引方法(force-directed method)^[3].近年来,也有人开始应用遗传算法来研究有向图画图算法^[4].在本文中,我们研究基于遗传算法的一般无向图的画图问题.一般无向图画图算法较少见.在文献[2]中,T. Kamada与S. Kawai将一个无向连通图看作是一个顶点为钢环、边为弹簧的动态系统,而画图过程则是一个逐步减少整个系统总能量的过程.在文献[3]中,T. M. J. Fruchterman和E. M. Reingold将一个无向连通图看作是一个力学系统,算法的最终目的是要使力学系统达到平衡.在以前的研究中,不同的无向图画图算法有不同的算法框架,所用的知识也有所不同,但一般根据应用的不同、个人美学观点的不同,有时需要对美观准则作出修改以适应不同的需要,而以前的画图算法由于美观准则嵌入算法之中,故难以修改,因而我们需要一种易于修改的画图算法以适应不同的需要.在本文中,我们将一般无向图的画图问题看成是一个函数优化问题,然后用遗传算法^[5]来求目标函数的最优解的近似值,从而得到无向图自动画图算法的一个一般框架.不同的画图算法的算法框架都一样,所不同的只是反映无向图画图问题的审美标准的目标函数.其优点在于,算法统一、方法简单、便于修改并且易于并行实现.而且以前的无向图画图算法都是用来画连通图的,对于画非连通图,则采用分别画每一个连通分支,然后将不同的连通分支拼接起来的方法,而我们的算法则可以直接用来画非连通图.

* 本文研究得到国家自然科学基金(No. 69635030)、国家863高科技项目基金(863-306-ZT06-06-3)和湖北省重大科技项目基金(No. 982P3119)资助.作者黄竞伟,1956年生,教授,主要研究领域为算法设计与分析,演化计算.康立山,1934年生,教授,博士生导师,主要研究领域为并行计算,演化计算,计算数学,计算机科学理论.陈毓屏,女,1942年生,副教授,主要研究领域为演化计算,并行计算,计算机软件.

本文通讯联系人:黄竞伟,武汉 430072,武汉大学软件工程国家重点实验室

本文 1998-09-25 收到原稿,1999-01-22 收到修改稿

1 画图问题

一个图 $G=(V, E)$ 是一个顶点与边的集合对, 其中 V 为顶点的集合, E 为边的集合. 图 G 的一种画法 θ 是将 G 的每个顶点 $v \in V$ 映射到平面上的一个点 $P(v)$, 并将 G 的每条边 (u, v) 映射成平面上的一条端点分别为 $P(u)$ 和 $P(v)$ 的简单约当曲线. 若 G 的每条边都被映射成一条直线段, 我们称画法 θ 是一个直线画法. 若 G 的每条边都被映射成一条折线, 称画法 θ 是一个折线画法. 若 G 的每条边都被映射成一条平行、垂直线段链, 称画法 θ 是一个正交画法. 若 G 的任两条边的映像都不相交, 称画法 θ 是一个平面画法.

在本文中, 我们研究无向图的直线画法, 这样, 画图算法的目的是将图画到某个有限区域中 (譬如显示屏上), 而算法的主要任务是在一定的美观准则下, 对图中的每一个顶点 v 指定一对坐标 $(x(v), y(v))$, 一旦每个顶点都被指定一对坐标后, 则将图画出来就是一件很容易的事情了, 这只需在有边相连的顶点之间画出一条直线段即可. 对于画一般的无向图, 根据应用和观点的不同, 有不同的美观准则. 但人们仍普遍接受以下美观准则:

(1) 尽可能将顶点均匀地分布在有限区域中;

(2) 尽可能避免边的交叉;

(3) 尽可能使边长一致;

(4) 反映图的内在对称性.

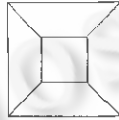


Fig. 1 Drawing without crossing edges

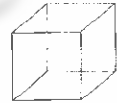


Fig. 2 Drawing with crossing edges

图1 无交叉边的画法

图2 有交叉边的画法

在我们的算法中, 只有准则

(1)、(3)得到了明显的体现, 但算法的结果也基本满足准则(4). 我们注意到, 准则(3)在有些情况下并非必需. 例如对立方体图, 我们有两种画法, 如图1和图2所示. 虽然图1中的画法没有边的交叉, 但图2比图1看起来要好一些.

2 无向图画图问题的遗传算法

遗传算法是近年来发展迅速的一种随机搜索算法, 它的一个重要应用领域是函数优化. 用遗传算法求解画图问题的方法是将画图问题转化为函数优化问题, 如何设计目标函数, 使之能够反映无向图画图问题的美观标准是用遗传算法求解画图问题的关键. 下面, 我们首先讨论如何设计目标函数.

对目标函数的设计基于下面两点观察:

(1) 任何两个顶点不应靠得太近,

(2) 有边相连的顶点应相互靠近.

我们用下面的函数作为目标函数.

$$f = \sum_{i=1}^n \sum_{j=i+1}^n \frac{L^2}{|p_i - p_j|} + \sum_{(v_i, v_j) \in E} \frac{|p_i - p_j|^2}{L} \quad (1)$$

其中 P_i 为顶点 V_i 在平面上的映像, $|p_i - p_j|$ 表示 p_i 与 p_j 之间的距离, $L = k\sqrt{\frac{S}{n}}$ 为两点间的理想边长, n 为图 G 的顶点个数, S 为有限区域的面积, k 为由实验所得的参数. 若条件(1)不满足, 则使式(1)中的第1个和式增加; 若条件(2)不满足, 则使式(1)中的第2个和式增加, 因而画图问题转化为求式(1)的最小值点.

若用 $x(v), y(v)$ 分别表示顶点 v 在平面上的映像 $P(v)$ 的 x 坐标和 y 坐标, 则因为我们希望将图画到有限区域 $a \leq x \leq b, c \leq y \leq d$ 中, 因而我们的目的是求解以下带约束条件的目标函数最小值问题.

$$\text{Min} \sum_{i=1}^n \sum_{j=i+1}^n \frac{L^2}{\sqrt{(x(v_i) - x(v_j))^2 + (y(v_i) - y(v_j))^2}} + \sum_{(v_i, v_j) \in E} \frac{(x(v_i) - x(v_j))^2 + (y(v_i) - y(v_j))^2}{L} \quad (2)$$

其中 $a \leq x(v_i) \leq b, c \leq y(v_i) \leq d, i=1, 2, \dots, n$.

用遗传算法求解式(2)的思想是: 首先用称为染色体的一种简单数据结构对式(2)可能的解 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 进行编码, 然后将一组遗传算子作用于一个解的编码群体, 即染色体的群体上, 使其能够保持一些关键信息, 通过不断演化, 最后得到式(2)的最小值点或近似最小值点.

设计一个遗传算法的基本步骤由编码、设计适应度函数、制定选择策略、选择控制参数、设计遗传算子以及给出终止准则 6 个部分组成. 下面分别对这 6 个步骤进行讨论.

(1) 编码

与传统的遗传算法不同,对式(2)中每个可能的解 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, 我们用一个长度为 $2n$ 的实数串 $(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$ 来表示, 这样做的一个主要好处是可以避免解码过程, 减少算法时间.

(2) 设计适应度函数

因为本问题是求最小值点,故我们可以选择一个常数 $C_{max} \geq f(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$, 直接从 $f(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$ 来设计适应度函数 $F(x_1, y_1, x_2, y_2, \dots, x_n, y_n) = C_{max} - f(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$.

(3) 选择策略

为了防止早熟发生,我们首先用 Sigma 比例变换技术对个体的适应值进行变换,即对个体 i 的适应值 $f(i)$, 首先用公式

$$ExpVal(i) = \begin{cases} 1 + (f(i) - f(t)) / 2\sigma(t), & \text{若 } \sigma(t) > 0 \\ 1, & \text{若 } \sigma(t) = 0 \end{cases}$$

将 $f(i)$ 变换到 $ExpVal(i)$, 其中 $f(t)$ 为第 t 代群体平均适应值, 而 $\sigma(t)$ 表示第 t 代群体的标准方差. 然后对 $ExpVal(i)$ 应用基于适应值比例的选择策略, 但保留适应值最高的染色体.

(4) 控制参数

在我们的算法中,取 $N=20, P_c=0.75, P_m=0.015$. 它们分别表示群体的大小、杂交概率与变异概率. 最大演化代数 T 一般为图的顶点数与边数的函数, 在实际应用中可根据图的顶点数与边数的大小取不同的值.

(5) 遗传算子设计

对于杂交算子,我们采用单点杂交,即若所选择的两个父体为

$$A = (a_1, a_2, \dots, a_{2n}), B = (b_1, b_2, \dots, b_{2n}),$$

选择杂交位置为 3, 则所得到的两个新个体分别为

$$A' = (a_1, a_2, a_3, b_4, \dots, b_{2n}); B' = (b_1, b_2, b_3, a_4, \dots, a_{2n}).$$

对于变异算子,我们采用如下的非一致变异算子:设父体为 $A = (a_1, a_2, \dots, a_k, \dots, a_{2n})$, 第 k 个分量被选出来进行变异. 假设第 k 个分量为 y 坐标, 其取值范围为 $[c, d]$, 则变异后所得到的新个体为

$$A' = (a_1, a_2, \dots, a_k', \dots, a_{2n}),$$

其中

$$a_k' = \begin{cases} a_k + \Delta(t, d - a_k), & \text{如果 } Random(2) = 0; \\ a_k - \Delta(t, a_k - c), & \text{如果 } Random(2) = 1. \end{cases}$$

这里, $Random(2)$ 产生 0 或 1 的一个随机整数, 而 $\Delta(t, y) = y \cdot (1 - r^{(t-T)^5})$, 其中 r 为 $[0, 1]$ 之间的一个随机数, t 为当前演化代数, T 为最大演化代数. 显然, $\Delta(t, y)$ 的值为 $[0, y]$, 且当 t 趋于 T 时, $\Delta(t, y)$ 趋于 0. 这表明, 在演化初期, 变异算子在较大的范围内搜索; 而在演化后期, 变异算子起到一种局部微调的作用.

(6) 终止准则

采用算法运行若干代后终止算法. 基于遗传算法的无向图的通用画图算法框架描述如下:

Procedure Graph_Drawing_By_GA;

Begin

$t := 0; initialize(P(t)); Evaluate(P(t));$

While $t < T$ Do

Begin

$P1 := Select(P(t)); P2 := Crossover(P1);$

$P(t+1) := Mutate(P2); Evaluate(P(t+1));$

$t := t + 1;$

End;

Let $(x_1, y_2), (x_2, y_2), \dots, (x_n, y_n)$ be the points with themaximum fitness among the population;

Draw_Graph(G, X, Y);

End.

其中 Draw_Graph(G, X, Y)过程画出顶点的 x 坐标为 $X = (x_1, x_2, \dots, x_n)$, 顶点的 y 坐标为 $Y = (y_1, y_2, \dots, y_n)$ 的图 G .

3 实验结果

我们已用 Delphi 编程实现上述算法. 对于一些简单无向图的画法, 我们得到了与文献[3]类似的结果. 但对于一些较复杂的无向图, 由于文献[3]中的算法对于图的某些初始状态容易陷入局部最小的近似值, 因而会得到较差的画法. 而我们的算法的稳定性较好, 实验表明, 画法的最终结果不依赖于图的初始状态. 而且我们的算法不仅可以用来画连通图, 也可以直接用来画非连通图, 只需将不同的连通分支画到不同的区域中即可. 图3是用我们的算法所画出的若干图例.

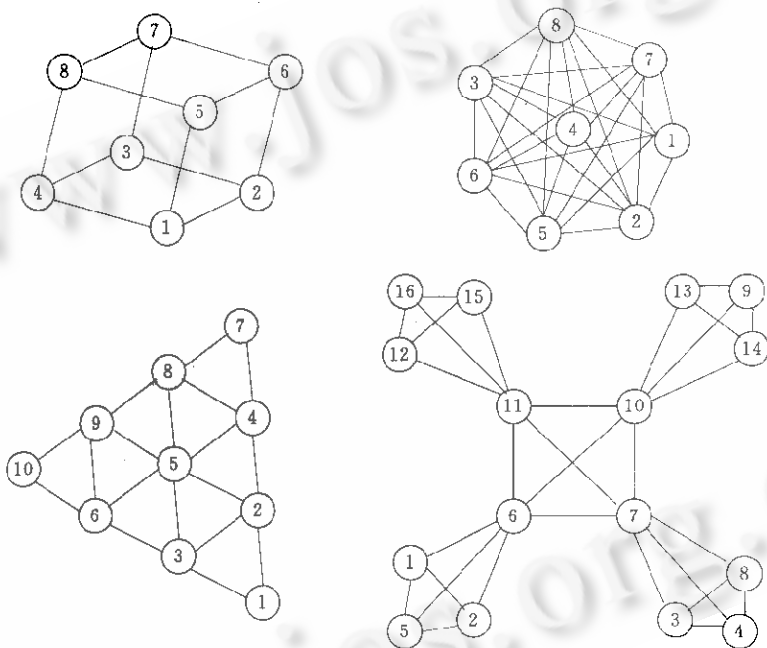


Fig. 3 Experiment results

图3 实验结果

4 结束语

本文用遗传算法求解画图问题. 在遗传算法的基础上, 我们得到了一个通用无向图画图算法框架, 不同的无向图画图算法的框架都一样, 所不同的只是所设计的反映无向图画图问题美学标准的目标函数. 例如, 若我们在上面的算法中用文献[2]中的能量函数代替目标函数, 则可得到与文献[2]类似的结果. 因而在基于遗传算法的通用无向图画图算法框架的基础上, 无向图画图算法的主要任务就是设计出能够反映各种美观标准的目标函数. 根据不同的应用以及不同的美观标准, 如何自动生成满足要求的目标函数是我们今后进一步的研究的方向.

参考文献

- 1 Battista G D, Eades P, Tamassia R *et al.* Algorithms for drawing graphs: an annotated bibliography. Computational Geometry: Theory and Applications, 1994, 4(5): 235~282

- 2 Kamada T, Kawai S. An algorithm for drawing general undirected graph. Information Letters, 1989,31(1):7~15
- 3 Fruchterman T M J, Reingold E M. Graph drawing by force directed placement. Software-Practice and Experience, 1991, 21(11):1129~1164
- 4 Kosak C, Marks J, Shieber S. Automating the layout of network diagrams with specified visual organization. IEEE Transactions on System, Man and Cybernetics, 1994,24(3):440~454
- 5 Michalewicz Z. Genetic Algorithms + Data Structures = Evolution Programs. 3rd edition, New York: Springer-Verlag, 1996

A New Graph Drawing Algorithm for Undirected Graphs

HUANG Jing-wei KANG Li-shan CHEN Yu-ping

(State Key Laboratory of Software Engineering Wuhan University Wuhan 430072)

Abstract In this paper, the authors transform the problem of undirected graph drawing to the problem of function optimization, then use genetic algorithms to find approximate optimal solutions of the objective function, and thus obtain a general structure of undirected graph drawing algorithms. The characters of the new method are: the structures of the different graph drawing algorithms are the same, the difference exists only in the objective functions which reflect aesthetic criteria. The advantages of the method are: unified algorithms, simplicity, easy modification and parallelism, and it can be used to draw non-connected graphs directly.

Key words Undirected graph, graph drawing, aesthetic criteria, algorithm, genetic algorithm.