

基于遗传算法的点状要素注记的整体最优配置

樊 红¹ 刘开军² 张祖勋¹

(1 武汉大学测绘遥感信息工程国家重点实验室, 武汉市珞喻路 129 号, 430079)

(2 华中科技大学管理学院, 武汉市珞喻路 1037 号, 430073)

摘要: 提出了一种点状要素自动注记的整体最优解的解决方案, 其核心算法采用具有全局搜索特性的遗传算法, 可以获取点状要素注记配置的(近似)全局最优解。

关键词: 遗传算法; 注记配置; 优化组合

中图法分类号: P231.5; P283.1

1 点状要素注记配置的基本规则和本文拟研究的问题

地图目标按其分布特点可分为点、线、面三种形式, 简单地考虑点状注记, 按照“所属关系清楚”、“避让性”和“习惯性”等总的原则, 结合我国地形图的图式规范, 在选用 1:25 万地形图居民地进行注记试验时, 考虑对其他专题图的通用性, 重点考虑以下三条原则。

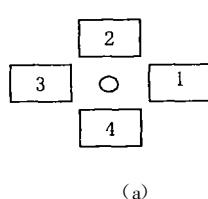
1) 备选位置及其优先级。点状要素备选位置通常有 4、5、8、n 备选位置等情况。4 备选位置如图 1(a)所示, 以正右为先, 其次分别是正上、正左、正下, 这些注记位置分别以优先级 1~4 来标记。5 备选位置如图 1(b)所示, 以正右为先, 其次分别是正上、正左、正下、正右竖排, 这些注记位置分别以优先级 1~5 来标记。8 备选位置如图 1(c)所示, 以正右为先, 其次分别是正上、正左、正下; 再次分别是右上、左上、左下、右下, 这些

注记位置分别以优先级 1~8 来标记。n 备选点位的情况如图 1(d)所示。

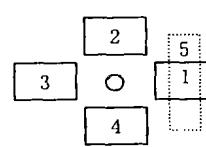
2) 禁止冲突。点状要素的注记之间不要相互压盖(称为冲突)。

3) 禁止和避免压盖。不压盖被注记要素和其他点状要素。避免压盖地图要素, 尤其应避开同种颜色的其他要素。如压盖不可避免, 选择压盖比较不重要的要素和不同颜色的其他要素。

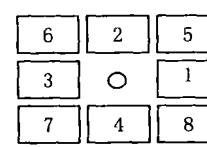
本论文研究主要考虑下列两类点状要素注记问题: ①对于地图上的点状要素, 采用 4 位置注记方案, 即分别选取点状要素的右、上、左、下的 4 个备选位置(如图 1(a)所示, 注记与点状居民地距离为 1mm), 然后进行优化处理, 使得冲突最少, 从而最终确定注记的位置。②对于地图上的点状要素, 采用 4 位置注记方案, 即分别选取点状要素的右、上、左、下的 4 个备选位置(如图 1(a)所示, 注记与点状居民地距离为 1mm), 然后进行优化处理, 使得冲突最少、对其他图层要素的压盖最小、注记位置最优, 从而最终确定注记的位



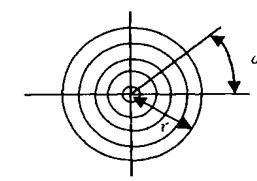
(a)



(b)



(c)



(d)

图 1 点状地物的备选注记位置

Fig. 1 Candidate Labeling Position of Point Feature

置。解决 4 个备选位置的算法容易推广到 5、8、n

备选位置的情况。

2 遗传算法解决点状要素注记问题方案

2.1 确定编码框架

地图自动注记是遗传算法的优化对象。地图的一个配置可以用一个向量表示，每个分量代表一个注记的定位。假定有 m 个备选位置，m 个备选位置可用 0~(m-1) 的编码表示，考虑 4 位置注记问题时，4 个备选位置可用 0~3 的编码表示。

采用整数向量对地图注记配置进行编码，一条染色体是一个向量，代表地图注记的一个配置方案。染色体的长度为 n (点状要素注记数)，分量(即基因)代表一个点状要素注记，基因的取值为 0~(m-1)，m 为备选位置数。对于 4 位置注记问题，基因的代码集为 {0, 1, 2, 3}。图 2 是一幅有 20 个点状要素的地图，图中也给出了它的一个配置及其对应的染色体编码。

2.2 产生初始种群

根据注记问题的特点，采取如下策略产生初始群体。

1) 随机产生一定规模的初始群体，为这些染色体随机选择每个基因位。

2) 根据注记问题的特点，对所有自由注记(取最优点位时不会发生冲突和压盖的注记)，在初始种群中，将所有染色体的相应的基因位直接取最优点位。

上述种群初始化策略将自由注记直接取最优点位，不再将该注记点位放在注记中进行处理。相当于减小了问题的规模，从而加速了算法的演化。

$$E_{\text{冲突}}(L_i) = \begin{cases} 1 & \text{如果 } \forall j, 0 < j < n, j \neq i, L_i \text{ 与 } L_j \text{ 不冲突} \\ 0 & \text{否则} \end{cases}$$

其中， L_i 表示第 i 个点状要素的注记。当注记无冲突时取值为 1，有冲突时取值为 0。这种适应度函数定义为地图上没有发生冲突的注记的总和。采用这种形式的适应度函数，遗传算法能较好地解决全局冲突。

$$E(L_i) = \begin{cases} W_{\text{压盖}} E_{\text{压盖}}(L_i, BF) + W_{\text{位置}} E_{\text{位置}}(L_i) & \text{如果 } (L_i) \text{ 无冲突} \\ 0 & \text{否则} \end{cases}$$

$$fit(L) = \sum_{i=1}^N E(L_i) \quad (3)$$

其中，取 $W_{\text{压盖}} = 100$ ； $W_{\text{位置}} = 1$ 。

式中， $E_{\text{压盖}}(L_i, BF)$ 为 L_i 的压盖评价值； BF 为背景要素的集合。采用简单压盖评价函数

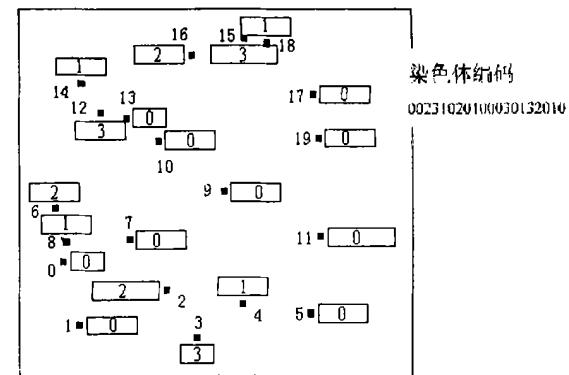


图 2 点状要素注记配置及其染色体编码

Fig. 2 Pointmap Labeling and Its Encoding String

2.3 确定适应度函数

注记问题的目标是搜索出质量最好的注记配置。自然地，适应度函数应定义为注记质量评价函数。为此，先定义注记质量评价的函数，它考虑冲突、压盖、位置优先级等因素，采用打分的方式，为每个注记的冲突、压盖、位置优先级等情况打分，然后将每个评分加权求和得到每个注记的总分，再对所有注记求和，得到整个注记配置的评分。评价值越高，注记质量越好，正好与对适应度表示的要求相一致（适应度值越大个体越优良）。根据这个思路，参照不同注记问题对优化目标的不同要求，可以定义相应的适应度函数。

2.3.1 使冲突最少的目标

先看点状注记问题①，它只考虑使冲突最少的优化目标。这时，定义适应度函数如下：

$$fit(L) = \sum_{i=1}^N E_{\text{冲突}}(L_i) \quad (1)$$

式中， $E_{\text{冲突}}(L_i)$ 为 0~1 冲突评价值函数。

$$E_{\text{冲突}}(L_i) = \begin{cases} 1 & \text{如果 } \forall j, 0 < j < n, j \neq i, L_i \text{ 与 } L_j \text{ 不冲突} \\ 0 & \text{否则} \end{cases} \quad (2)$$

2.3.2 使冲突最少、压盖最少和位置最优的目标

再看点状注记问题②，它有 3 个优化目标，适应度函数考虑冲突、压盖和位置优先级问题，因而适应度函数可定义为注记的下列评价函数：

时，它定义为被某注记所压盖要素中的最高重要性权重，无压盖时，评分为最高分 99 分，被压盖要素的重要性等级越高，则认为压盖越严重，压盖评分越低。令 BF_j 表示被注记压盖的第 j 个要素；

$Overlap(O_1, O_2)$ 表示两个对象 O_1, O_2 重叠。定义如下。令 OBF 表示被注记压盖的背景要素。

$$E_{\text{压盖}}(L_i, BF) = \begin{cases} 99 & \text{如果 } OBF \text{ 为 } 0 \\ 99 - \max\{W(OBF_0), W(OBF_1), \dots, W(OBF_{m-1})\} & \text{否则} \end{cases}$$

式中, $W(BF_j)$ 为背景要素所定义的重要性评价函数(其值称为重要性等级或重要性权重)。同样采用 0~99(取值为整数)的评分体系。对禁止压盖的要素评分为 99, 重要性程度越低, 评分越低, 最不重要的要素评分为 0。所以有:

$$W(BF_j) = \begin{cases} 0 & \text{重要性最低} \\ 1 \sim 98 & \text{重要性居中} \\ 99 & \text{重要性最高} \end{cases} \quad (5)$$

$E_{\text{位置}}(L_i)$ 为 L_i 所取备选位置的位置评价值。采用排序位置评价函数, 利用式(6)计算。当注记的备选位置有限多个, 而且数目不大可以枚举出来时, 例如点的 4 或 8 位置注记模式。将其按优先级由高到低的顺序排列, 令 $Pos(L_i)$ 表示 L_i 所取注记位置, $Order(Pos(L_i))$ 表示所取位置排序后的序号, 可将位置评价函数定义为 99 与位置序号之差。即优先级最高的位置评分最高为 99, 其余位置依序递减。于是有:

$$E_{\text{位置}}(L_i) = 99 - Order(Pos(L_i)) \quad (6)$$

采取这种适应度函数, 遗传算法在解决冲突的同时, 兼顾压盖和位置优先级的优化问题。

特别地, 如果只考虑使冲突最少和位置最优的目标, 则令式(3)中的 $W_{\text{压盖}}=0$, 得到:

$$E(L_i) = \begin{cases} W_{\text{位置}} E_{\text{位置}}(L_i) & \text{如果 } L_i \text{ 无冲突} \\ 0 & \text{否则} \end{cases}$$

$$fit(L) = \sum_{i=1}^N E(L_i) \quad (7)$$

其中, 取 $W_{\text{位置}}=1$ 。

考虑图 2 的例子, 图上给出了它的染色体表示。考虑使冲突最少和位置最优的目标, 采用适应度函数(7), 可计算它的适应度值为:

$$fit(L) = \sum_{i=1}^{20} E(L_i) =$$

父串 1	0	1	2	3	0	1	0	2
交叉点								
父串 2	1	0	1	1	2	2	1	0

父串 1	0	1	2	3	0	1	0	2
交叉点 1								
父串 2	1	0	1	1	2	2	1	0

素, 设 OBF 的总数为 m , OBF_j 表示其中的第 j 个背景要素。其含义如下:

$$\begin{array}{ll} 99 & \text{无压盖} \\ 99 - \max\{W(OBF_0), W(OBF_1), \dots, W(OBF_{m-1})\} & \text{有压盖} \end{array} \quad (4)$$

$$99 \times 10 + 98 \times 4 + 97 \times 3 = 1961.0$$

2.4 设计遗传算子

2.4.1 选择算子

注记算法中采取转轮式选择法作为选择方法。当种群规模较大时, 可选择使用精英策略, 即将往代的最优个体保留在选择后的群体中。转轮式选择策略实现的算法如下。

1) 计算个体的适应值 $fit(V_i)$ ($i=1, 2, \dots, n$);

2) 计算个体的累计适应度值 $Accfit(V_i)$ ($i=1, 2, \dots, n$) 和相对累计适应值 $RelAccfit(V_i)$ ($i=1, 2, \dots, n$)。

3) 生成一个 [0, 1] 内的随机数 r , 此处假设 $RelAccfit(V_0)=0$ 。如果 $RelAccfit(V_{i-1} < r \leq RelAccfit(V_i))$ ($i=1, 2, \dots, n$), 那么选择个体 i 。

2.4.2 交叉算子

整数向量编码常用的两种交叉算子为点式交叉和均匀交叉。试验表明, 这两种交叉算子对注记算法性能影响差别不大, 所以注记算法中采取点式交叉策略。

点式交叉算子又分为单点式交叉和多点式交叉。单点式交叉随机地在两个父串上选择一个交叉点, 然后交换这两个串对应的子串, 多点交叉则是一次随机生成多个交叉点, 然后间断交换父串的对应子串^[2]。图 3 给出了串长为 8, 基因代码集为 {0, 1, 2, 3} 的两个注记配置染色体点式交叉的例子。

2.4.3 冲突位变异

对于整数向量编码而言, 常用的变异有点式变异和均匀变异。前者选择单个点位, 后者根据一个模板选择点位, 对所选点位进行随机重定位。



图 3 点式交叉的例子

Fig. 3 Example of Pointal Crossover

对于地图注记, 笔者提出了一种新的变异算子, 称之为冲突位变异, 来代替常规点式变异和均匀变异。冲突点位变异的基本方法是选择冲突注记基因位, 随机生成一个注记点位编码, 代替原有基因。

试验证明, 冲突基因变异很有效。图 4 给出了在采用相同遗传参数(遗传参数为迭代数=300、种群=50、交叉率=0.75、变异率=0.2)情况下采用冲突位变异和均匀变异结果的比较。从图中可以看出, 使用冲突位变异的效果要明显好于均匀变异。主要有以下原因。

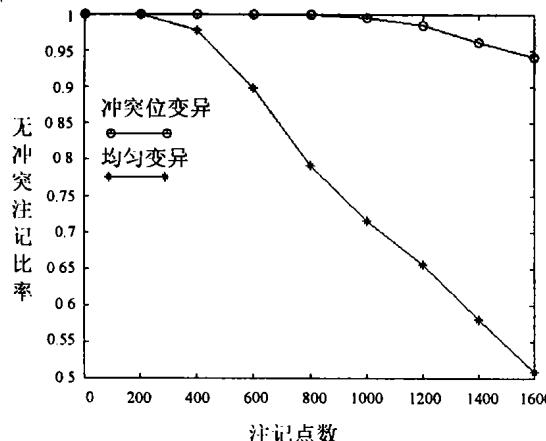


图 4 采用冲突位变异和均匀变异的结果的比较

Fig. 4 Comparison Between Mutate on Conflict Gene and Uniform Mutate

1) 点式变异和均匀变异是盲目的, 冲突位变异方法利用了注记冲突的启发式信息, 改进坏的子解、得到好的子解的可能性更大。

2) 在实际的地图中, 注记点稀疏的区域冲突

较少, 变异概率应小些, 而对于注记点密集的区域, 冲突点很多, 变异概率应比较大。冲突位变异适应了这种要求。

3 遗传算法试验结果及对比试验结果

3.1 遗传算法试验及其结果

对遗传算法方案进行的试验由模拟生成的随机图数据和实际的地形图数据两部分数据组成。它们包括: ①随机点状地图若干幅, 利用随机点状图生成算法产生; ②国家 1:25 万基础地形图数据库的全要素地形图 3 幅, 每幅地形图分为水系、道路、植被、境界等共 19 层。这 3 幅地形图分别包含点状居民地 2 511 个、1 651 个及 2 734 个。

从 50~3 000 点的点状要素地图遗传算法运行的演化试验可以看出, 遗传算法在迭代过程中, 最优个体适应度值逐渐增大, 达到一定代数后趋于稳定, 说明种群已经成熟。算法应该在种群成熟后终止。点数在 3 000 点内的点状要素地图一般在 300 代内成熟。

试验还包括不同复杂性的地图运用遗传算法情况的比较和同一地图使用不同参数的遗传算法情况的比较。试验表明, 相同遗传参数处理不同复杂程度问题时, 会得到不同的结果。在保持一定的种群规模时, 增加演化代数会增加解的适应度。但达到一定程度时, 种群成熟, 解不再改进。演化代数太少时, 简单地增加种群规模对解的改进不大。



图 5 H4810(2 511 点)遗传算法注记结果图

Fig. 5 Labeling Result of H4810 Applying Genetic Algorithm

试验中包括对3幅实际地形图的试验,所用3幅地形图属中等难度的问题,遗传算法在20~30s内(不包括预处理时间)能消除几乎全部冲突(所余冲突为1,0和2)。图5给出了H4810(2511点)注记的结果图。

3.2 比较试验结果及结论

为了验证本章提出的遗传算法的性能,将遗传算法与简单随机算法、爬山法、模拟退火法和神经网络算法进行了性能的比较。

采用的试验数据集为一组随机生成的点状地图。试验方法基于同样的数据背景和辅助数据条件进行,在使用算法前建立了同样的冲突查找表和压盖查找表。

试验主要对算法生成结果的质量进行了比较。为了便于进行试验比较,统一考虑点注记模式的4个位置注记问题,考虑下列两个优化目标:
①考虑冲突,不考虑位置优先级,不考虑压盖;
②考虑冲突,考虑位置优先级,不考虑压盖。

在只考虑冲突的情况下,用注记结果中的无冲突注记与注记总数的比率来表示注记质量。分别生成了200,400,600,800,1000,1200,1400,1600的8×5幅随机生成的点状要素图,图7给出了这些算法性能的比较结果。图上显示的结果是5幅随机地图注记结果的一个平均值。从上述的比较试验可以得到以下两个结论。

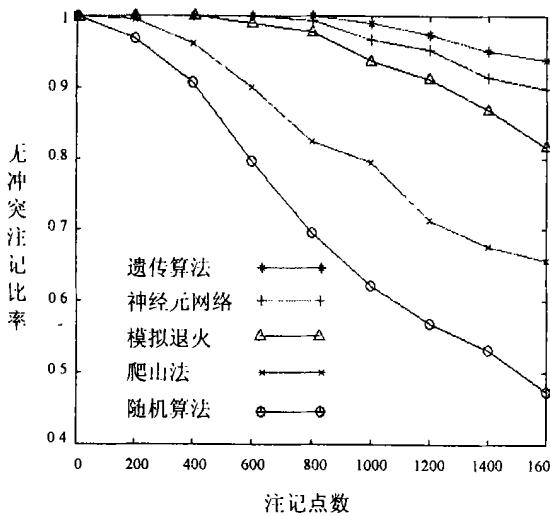


图6 5种算法性能比较

Fig. 6 Performance Comparison Among Five Algorithms

1) 从图6可以看出,对于相同复杂情况的图件,遗传算法的解的质量最高;其次是神经元网络方法;最后才是模拟退火算法、爬山法,随机算法质量最差,它的注记质量是算法能得到的解的质量的下限。从注记质量来看,遗传算法>神经元

网络算法>模拟退火>爬山法。从综合性能而言,遗传算法最好。

2) 本文提出的遗传算法是一种强壮、易扩展、性能优良的自动注记算法,它有下列优点:易于加进对其他优化因素的考虑,扩展性好;编码表可以由问题确定,易于根据问题扩展;GA十分强壮,不会产生非法解;GA的参数容易调制;GA的主要参数已由系统确定,参数调制工作量很小。

4 结语

为了研究的方便,在本文的研究中对复杂的地图注记情况进行了简化,只考虑了冲突最少、对其他图层要素的压盖最小、注记位置最优等目标,而实际应用中,地名注记所面临的情况要复杂得多,不同的地物类型、不同的比例尺、不同的专题图,对注记的要求不同。通过修改适应度函数,使之兼顾更多的因素,可以进一步完善现有的算法,使之更为灵活使用,能处理更为复杂的注记问题。同时,遗传算法解决线状要素和面状要素的方案,也要比遗传算法解决点状要素的方案复杂,这也是需要进一步研究的问题。

参 考 文 献

- Anthony C C, Christopher B J. A Prolog Rule-based System for Cartographic Name Placement. Computer Graphics Forum, 1990, 9(2): 109~126
- Christopher B J, Anthony C C. Rule-based Name Placement with Prolog. Auto-Carto 9, Baltimore, Maryland, 1989
- Cormely R G. An LP Relaxation Procedure for Annotating Point Feature Using Interactive Graphics. Auto-Carto 7, Virginia, 1985
- David S J, Umit Basoglu. The Use of Artificial Intelligence in the Automated Placement of Cartographic Names. Auto-Carto 9, Baltimore, Maryland, 1989
- Doerschler J, Freeman H. An Expert System for Dense Map Name Placement. Auto-Carto 9, Baltimore, Maryland, 1989
- Freeman H, Ahn J. Automap: An Expert System for Automatic Map Name Placement. The First International Symposium on Spatial data Handling, 1984
- Gail E L, Thomas K P. Integration of Name Selection and Name Placement. Auto-Carto 8, Virginia, 1986
- Freeman H, Ahn J. On the Problem of Placing Names in a Geographic Map. Int. J. of Pattern Recog. and Art. Intel., 1987, 1(1): 121~140
- Hiersch S A. An Algorithm for Automatic Name Place-

- ment Around Point Data. *The American Cartographer*, 1982, 9(1): 5~17
- 10 Imhof E. Positioning Names on Maps. *The American Cartographer*, 1975, 2(2): 128~144
- 11 James E M. Automated Feature and Name Placement On Parallel Computers. *Cartography and Geographic Information Systems*, 1993, 20(2): 69~82
- 12 Jeffrey S D, Freeman H. A Rule-based System for Dense-map Name Placement. *Communications of the ACM*, 1992, 35: 68~79
- 13 Ahn J, Freeman H. A Program for Automatic Name Placement. *Auto-Carto*, Ottawa, 1983
- 14 Christensen J, Marks J, Shieber S. An Empirical Study of Algorithms for Point-feature Label Placement. *ACM Transactions on Graphics*, 1995, 14(3): 203~232
- 15 Ebinger L R, Goulette M. Noninteractive Automated Names Placement for the 1990 Decennial Census. *Cartography and Geographic Information Systems*, 1990, 17(1): 69~78
- 16 Zoraster S. Integer Programming Applied to the Map Label Placement Problem. *Cartographica*, 1986, 23(3): 16~27
- 17 Yoeli P. The logic of Automated Map Lettering. *The Cartographical Journal*, 1972, 9(2): 99~108
- 18 Zoraster S. Expert Systems and the Map Label Placement Problem. *Cartographica*, 1991, 28(1): 1~9
- 19 樊 红, 张祖勋, 杜道生, 等. 基于神经网络模型求取注记配置最优解. *武汉测绘科技大学学报*, 1998, 23(1): 32~35
- 20 米凯利维茨. 演化程序. 北京: 科学出版社, 2000
- 21 潘正君, 康立山, 陈 屏. 演化计算. 北京: 清华大学出版社, 1998
- 22 周 明, 孙树栋. 遗传算法原理及应用. 北京: 国防工业出版社, 1999

作者简介: 樊红, 副教授, 博士。主要研究方向为地图学与地理信息系统理论与应用、空间影像信息处理。代表成果有基于最优解理论的地图注记自动化与智能化研究、基于航空影像的城市变化检测。发表论文 20 余篇。

E-mail: fh@hp01.wtusm.edu.cn

A Robust Genetic Algorithm for Automated Map Name Placement

FAN Hong¹ LIU Kajun² ZHANG Zuxun¹

(1 National Laboratory for Information Engineering of Surveying Mapping and Remote Sensing,
Wuhan University, 129 Luoyu Road, Wuhan, China 430079)

(2 Management School, Huazhong Science and Technology University, 1037 Luoyu Road Wuhan, China 430073)

Abstract: The traditional algorithm for automated map name placement and its disadvantage are firstly analyzed in this paper. A new global optimization algorithms that called genetic algorithm is put forward to solve the point-labeling problem. According to the properties of map labeling, the basic design schemes and strategies of applying genetic algorithm to solve the map name placement is detailed in this paper. First, an integer vector coding scheme is adopted in the algorithm, which uses an array (or string) of integers to represent a configuration of map labeling. The initial strings can be obtained randomly by generating a random integer in a specified scope for every element (gene) of all strings. Then a fitness function of map labeling is represented as a quality evaluation function of map labeling.

Lastly some typical experiments are elaborated and some results obtained by the automated map-labeling program based on genetic algorithm are presented. In the meantime, a comprehensive experiment is conducted to compare this algorithm with climbing algorithm, annealing algorithm and Hopfield neural network method, and experimental results have shown that the performance of genetic algorithm is superior to those of other several traditional algorithms.

Key words: genetic algorithm; map name placement; optimal combination problem

About the author: FAN Hong, associate professor, Ph.D. Her research interests are in the theory and application of cartography and GIS and spatial image processing. Her representative achievements include the research on automatic map labeling and the research on change detection of the urban areas based on aerial images. Her published papers are more than twenty.

E-mail: fh@hp01.wtusm.edu.cn