

基于电荷排斥模拟法的均匀采样布局*

陈柏松¹ 潘瑜春² 王纪华² 付卓³ 曾志炫²

(1. 北京师范大学地理学与遥感科学学院, 北京 100875; 2. 国家农业信息化工程技术研究中心, 北京 100097;
3. 环境保护部卫星环境应用中心, 北京 100094)

【摘要】 通过模拟理想同种点电荷排斥运动的方法实现 N 个采样点均匀分布在给定的带有限制区的不规则多边形采样区域内。首先将多边形采样区域以及内部限制区域的各条边离散化为一系列位置固定的点电荷, 然后在多边形内随机放置 N 个点电荷, 并通过逐个“放抓”点电荷并让它们受电场力运动, 并最终达到合力与速度都接近零的稳定状态, 此时 N 个点电荷的位置就均匀分布在指定区域内, 可以作为 N 个样点的均匀采样布局。分析与实验结果表明, 这种方法是一种在复杂多边形区域内实现采样点均匀分布的简单而高效的方法。

关键词: 均匀采样布局 库伦力 电荷排斥 模拟

中图分类号: S159-3 **文献标识码:** A **文章编号:** 1000-1298(2011)09-0181-05

Implementation of Even Sampling Layout by Electrons Repulsion Simulation

Chen Baisong¹ Pan Yuchun² Wang Jihua² Fu Zhuo³ Zeng Zhixuan²

(1. School of Geography, Beijing Normal University, Beijing 100875, China
2. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China
3. Satellite Environmental Application Center, Ministry of Environmental Protection of the People's Republic of China, Beijing 100094, China)

Abstract

A method for an even distribution of N sampling points within a given polygon which contained inaccessible constraints was presented based on the simulation of ideal electrons' repulsion movement. First, the edges of the polygon and the inner constraints into bunches of fixed electrons were described. N electrons can then be put into the polygon in random positions. And then they were “free and hold” one by one and were moving under Coulombic force. Finally, a balance was achieved with their resultant force and velocities approach to zero. At this time, these N electrons were evenly distributed within the given region; thus, their layout was considered to be an even sampling layout of N samples. Theoretical analysis and the experimental results indicated that electrons' repulsion simulation is a simple and efficient method for implementing even sampling layout of N samples within a complex polygon.

Key words Evenness sampling layout, Coulombic force, Electrons' repulsion, Simulation

引言

在农业环境监测及土壤属性制图等工作中, 采样点数量与布局直接影响到采样的代表性, 并最终影响结果准确性。通常情况下, 采样点数往往受限于时间和经济成本, 在一定采样点数下, 如何布局采

样点更好, 是需要研究的问题。系统采样能使采样得到的数据更符合整个区域的情况, 或插值后更接近整个区域的目标变量的实际分布^[1], 因此, 该采样布局方法被视为在采样区无任何目标变量先验知识且采样点数相对较少情况下的最好采样布局^[2], 并在采样领域中被广泛采用。系统采样的

收稿日期: 2010-11-17 修回日期: 2010-12-29

* 国家自然科学基金资助项目(40971237)

作者简介: 陈柏松, 博士生, 主要从事 GIS 理论与应用研究, E-mail: baisongchen@126.com

通讯作者: 王纪华, 研究员, 博士生导师, 主要从事农业信息化与 3S 技术研究, E-mail: wangjh@nrcita.org.cn

目标是使采样点在空间上均匀分布,这个问题可以抽象为“如何将 N 个点均匀地散布在一个二维多边形内”的几何问题求解。目前已有不少学者对此进行了研究,从早期的三角形格网、矩形格网、蜂窝形格网等规则多边形格网法^[3~6],发展到最近的利用模拟退火算法结合一定准则或目标函数来实现样点均匀分布的方法^[7~9]。多边形格网的方法比较简单,但实际采样区域往往不是规则区域,而且里面还可能不存在不能采样的限制区域,所以在实际采样中,多边形格网法很难实现给定样点数的均匀分布,而且由于边界效应的存在,该方法已被证明不是一种最优的均匀采样布局^[7]。利用模拟退火算法实现的采样布局优化,理论上是可以实现 N 个点在任意平面多边形内均匀分布,但由于模拟退火冷却表参数设置的问题直接影响到算法的全局收敛性,以及描述均匀分布的目标函数本身的复杂性,使得这种方法需要大量耗时的计算,且实际得出的布局并不如预想的均匀。基于此,本文提出一种基于电荷排斥模拟的采样点布局方法。

1 电荷排斥模拟方法

1.1 物理学基础与假设

2 个理想点电荷在电场中受到的库伦力为

$$F = \frac{kQ_1Q_2}{r^2} \tag{1}$$

式中 F ——库伦力
 k ——库伦力常数
 $Q_1、Q_2$ ——两个点电荷的带电量
 r ——两个点电荷的距离

为了计算简便,可以设 k 为 1,电场内(即采样区域的多边形内)的 N 个电荷的带电量均为 1,而电场边(即采样区域的多边形各边上)固定的电荷电量均为 0.1 或可变量。

根据向量相加的物理意义,单个场内点电荷所受合力为

$$\boldsymbol{F} = \sum \boldsymbol{F}_i \tag{2}$$

式中 $\boldsymbol{F}$ ——某场内电荷 A 所受的合力向量
 $\boldsymbol{F}_i$ ——除去电荷 A 本身外的其他某个点电荷(包括场内和边上)对 A 作用力向量,大小可由式(1)计算,方向是某个电荷指向电荷 A

假设某个很短的时间间隔 t 内,某个点电荷 A 受到库伦力的合力可视为恒定不变的,由此合力产生的加速度为

$$a = F/m \tag{3}$$

式中 a ——电荷 A 在受合力大小为 F 时的加速度,其方向与合力的方向一致

m ——点电荷 A 质量,设为 1

假设某个很短的时间间隔 t 内,某个点电荷 A 在静止状态下由加速度 a 的作用产生的位移大小为

$$S_A = \frac{1}{2}(v_0 + v_t)t = \frac{1}{2}(0 + at)t = \frac{1}{2}at^2 \tag{4}$$

其中 $v_t = at$

式中 S_A ——位移,方向与加速度的方向,也就是合力的方向一致

v_t ——电荷 A 由静止开始经过很短的时间间隔 t 后的速度

1.2 电荷排斥模拟

为了方便模拟,本文假设理想同种点电荷的运动是在光滑的平面上进行的。为了避免点电荷逸出采样区域的边界,或者进入采样区域内的限制区域内,首先要将采样区域及其内部限制区域的边界进行离散化,即由某边的起点开始按较小的距离间隔 D 将带电量为 Q_2 的点电荷逐个布设并固定在该边上,用同样的方法依次将各条边离散化使其成为线状排列的点电荷序列。离散化完成后往电场中(即在采样区域多边形内但又不限制区域多边形内的位置)按随机位置添加 N 个带电量均为 Q_1 的点电荷。假设这 N 个电荷随机放进场内时,都被固定在它们的各自的初始位置上,然后释放第 1 个电荷,则它立即就会受除它本身外的其他电荷(包括边界上和电场内的电荷)的排斥合力(式(1)、(2))作用而由静止开始运动,假设它所受合力在很短的时间间隔 t 内保持恒定,则可以计算它经过 t 时刻运动后的位置(式(3)、(4))。此时又将这个电荷“抓住”让其静止在新的位置,然后释放第 2 个电荷,同样地让其受力运动相同的时间间隔 t ,又同样地“抓住”它,依次释放各个电荷,让它们都单独地运动时间间隔 t ,直到场内的第 1 个到第 N 个电荷都被“放抓”一遍,称作一次循环。循环一直进行下去直到某次循环中所有场内可移动的 N 个点电荷中受合力最大的点电荷的所受的合力接近零,并且运动速度也接近零,即达到了稳定状态,此时这 N 个点电荷就均匀地分布在电场内部,它们的位置布局即可作为 N 个样点均匀地分布在给定采样区域内的采样布局。

1.3 算法流程

根据上述电荷排斥运动的规律,电荷排斥模拟的实现流程如图 1 所示。可以将整个过程总结为如下主要步骤:①离散各边。②随机放入 N 个电荷。③逐个“放抓”电荷。④达到均匀稳定状态。⑤输出 N 个电荷布局。

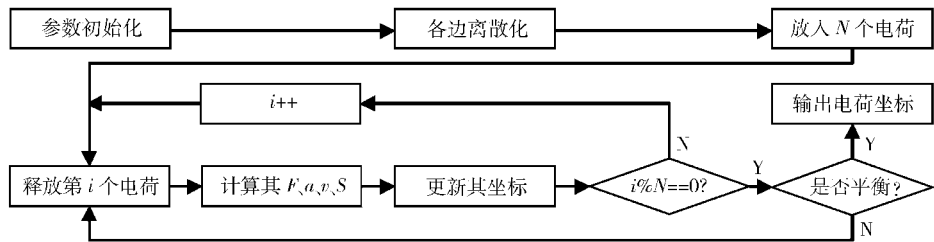


图 1 电荷排斥模拟的实现流程图

Fig. 1 Flowchart for implementing electrons' repulsion simulation

在程序实现过程中需要注意的是：①边上的电荷始终保持静止，它们从离散化开始就一直被固定在原来的位置，它们的作用是产生向内的排斥力防止场内的 N 个电荷逃逸出去，边界离散化时放置在边界上的电荷的间隔要适中，应该保证最短边上至少要有 10 个以上的电荷存在，以防止场内电荷逸出，另外边界电荷带电量与场内电荷带电量大小的比例要适中，这个比例影响到边界距离，边界电荷带电量太小可能使得场内电荷过度靠近边界，甚至逸出边界。②往电场中随机放置 N 个点电荷的位置不能相同或者太接近，以免在求合力时出现除数为零的问题。③刚开始随机放置的电荷不应太靠近边界，以免碰巧其附近有很多聚集的电荷将其“推”出边界外或者“挤”进约束区域内。④电荷每次受力运动短暂时间间隔 t 后，被“抓住”可使其在

下次循环中再释放时它的初速度为零，从而大大降低运动的复杂性而简化计算，因为如果不能抓住，下次循环到该电荷的时候该电荷会有初速度，这个初速度与新的合力方向很可能不相同，所以又需要进行分解计算。

2 结果与讨论

2.1 简单矩形区域的电荷排斥模拟结果

图 2 是电荷排斥模拟平衡后得到的均匀采样布局图，边界是 100×100 的正方形，没有限制区域，电场中电荷的电量设置为 1，离散化边上的电荷电量设置为 0.1；图 2a 中 $N = 16$ ，电荷排斥模拟正好得到 4×4 的样点排列格局，图 2b、2c、2d 中电荷数量逐渐增大，从图形可见，它们得到的均匀布局也是比较理想的，这些结果证明了该方法的有效性。

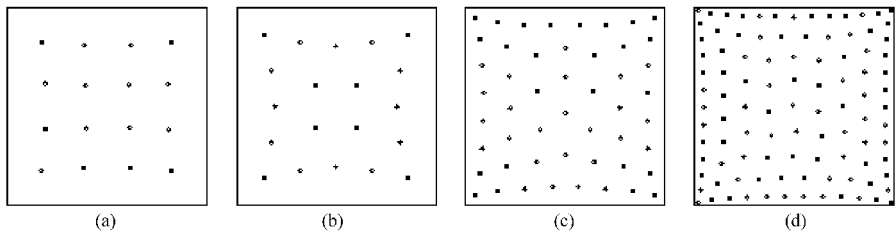


图 2 不同数量电荷在 100×100 方形区域中的均匀稳定布局图

Fig. 2 Even and stable layouts for $N = 16, 20, 50$ and 100 electrons in an 100×100 square

(a) $N = 16$ (b) $N = 20$ (c) $N = 50$ (d) $N = 100$

2.2 带限制区域的多边形电荷排斥模拟结果

图 3 是电荷在一个带有 3 个限制区域的多边形内的排斥模拟平衡后得到的均匀采样布局图。同样可见，各个布局中的电荷都是均匀地分布在多边形内，但是却不在多边形中的限制区域内。这一点

在实际应用中比较重要，因为在采样点布局设计时很可能碰到限制区域的情况，例如要对某地块或农场进行土壤水分或有机质含量制图，但是采样研究的区域中存在一口鱼塘或较大的建筑物等自然限制的因素使得采样在这些区域内无法进行，因此在考

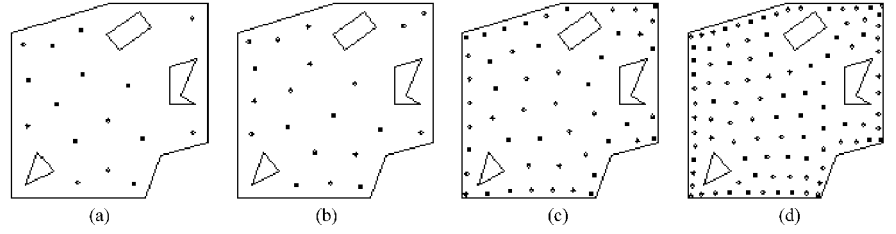


图 3 不同数量电荷在带限制区域的多边形区域中的均匀稳定布局图

Fig. 3 Even and stable layouts for $N = 16, 20, 50$ and 100 electrons in a polygon with 3 constraints

(a) $N = 16$ (b) $N = 20$ (c) $N = 50$ (d) $N = 100$

考虑采样点均匀分布的时候同时应该考虑不让采样点落入这些区域。

2.3 均匀度、网型与效率比较

图 4 是 3 种不同的方法实现的 23 个采样点在一个 100 × 100 方形区域中均匀分布的采样布局,其中图 4a、4b 引自文献[7]。从图形看,这几种方法产生的采样布局都比较均匀,难以量化比较哪个更好,所以需要构建一个均匀度指标用于量化衡量不同布局的均匀度。在一维空间中, N 个点均匀地分布在一维空间中时它们的距离间隔相等,也就是每个点占有其周围一维空间中的长度相等;类似地,在二维空间中如果每个点占有其周围的二维空间的面积相同,则可以认为这些点均匀地分布在这个空间中。因此,对于 N 个点均匀地分布在某一采样区域内,可以构建这些点的泰森多边形网,然后,计算它们的泰森多边形面积的变化积累来构建一个点在平面内分布均匀度指标

$$E = \frac{A}{\sum_{i=1}^N |S_i - S|} \tag{5}$$

其中 $S = \frac{A}{N}$

式中 N ——采样点或电荷的个数
 S_i ——第 i 个点对应的泰森多边形面积
 S ——名义泰森多边形面积,即将整个采样区域的面积 A 平均分给 N 个点,每个点所应得的泰森多边形面积

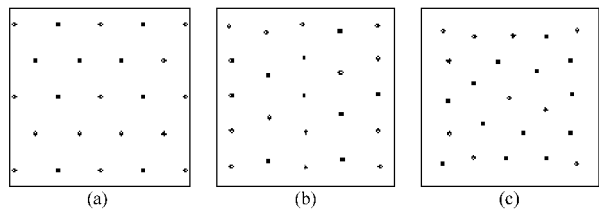


图 4 3 种方法实现 $N = 23$ 在 100×100 区域中均匀分布的采样布局

Fig. 4 $N = 23$ implementations from three methods in an 100×100 square

(a) 正三角形网 ($E = 6.00$) (b) 模拟退火与 MMSD 准则 ($E = 9.63$) (c) 电荷排斥模拟 ($E = 9.98$)

这个指标反映了点在平面空间内分布的均匀度,在同一区域内分布 N 个采样点,不同的布局可能带来不同的均匀度,它们的均匀度就可以用这个无量纲的 E 来衡量, E 值越高说明其对应的布局越均匀。由图 4 可见,规则三角形网的均匀度最小,而模拟退火法产生的采样布局的均匀度要略小于电荷排斥模拟法。

针对采样问题,图 4a 有部分采样点离采样区域

的边界太近,根据空间自相关原理,位置相近的两点具有的属性值会越接近,故样点无需太靠近边界就能反映边界上某点的属性值,所以图 4a 中样点过度靠近边界会造成其“捕捉属性信息能力”的浪费,再者这种指定 N 个点的等边三角形网很难在一个带有限制区域的复杂多边形内产生,至今尚无有效的方法。而图 4b 是使用模拟退火与 MMSD 目标函数得到的均匀采样布局,由于模拟退火算法冷却表参数设置的复杂性以及其对全局最优的重大影响^[10],在实际应用中很难得到一个最优的均匀采样布局,而往往只能得到一个大致均匀的采样布局。还由于模拟退火算法在迭代过程中的随机性,使得这种方法的计算量非常大,而且这种方法在实现过程中涉及繁琐的初始化过程以及迭代中复杂的目标函数计算,更大幅增加了此方法的计算量。图 4c 是使用本文提出的电荷排斥模拟的方法计算得到的均匀采样布局,其中靠近边界的电荷分布呈现弧形的特点,即靠近边界中点的地方的电荷离边界越远,这可能是由于电场等势面造成的,但并不影响点在整个区域的均匀分布。

利用模拟退火算法结合 MMSD 准则及电荷排斥模拟方法在 N 个点均匀采样布局的实现中每次迭代计算的时间频度分别是 $T(GN)$ 及 $T(E + N)$,其中 G 是一个覆盖采样整个采样区的非常精密的网格的点数,所以 G 的数值往往能达到万级^[7,11],而 E 则是电荷排斥模拟用于离散各边的电荷数, E 的数值通常是百级。从每次迭代的时间频度可见,电荷排斥模拟方法具有更高的效率。

分别用模拟退火与电荷排斥模拟两种方法实现图 4b、4c 的均匀采样布局 100 次所需的平均时间如表 1 所示。

表 1 两种方法实现图 4b、4c 的均匀采样布局所需的平均时间			
Tab. 1 Mean time using in the implementation of the even sampling layouts of fig. 4b and 4c s			
方法	初始化	迭代	平均总时间
模拟退火 + MMSD	0.9	1 620.7	1 621.6
电荷排斥模拟	0.4	46.2	46.6

注:机器配置:CPU 奔腾 D3.2 GHz,2 GB 内存,WinXP 系统,两个程序均由 C++实现,数据可能会随程序参数设置变化而变化。

可见电荷排斥模拟的方法实现均匀采样布局要比使用模拟退火与 MMSD 目标函数的方法原理上要简单,运行效率上也高很多。

3 结论

(1) 提出通过模拟理想同种点电荷排斥运动行

为,实现了 N 个点均匀分布在给定二维平面多边形区域内的方法,这种均匀样点布局也是 N 个点均匀分布在给定多边形内的计算几何问题的一种解法。

(2) 电荷排斥模拟方法相比规则多边形格网方法更适用于在带有限制区域采样区的均匀采样布局的实现,且效率比模拟退火方法高很多,因此它比较

适合于带有限制区域的多边形采样区域内实现 N 个样点的均匀分布问题。

(3) 边界电荷带电量与场内电荷带电量大小的比例(Q_2/Q_1)要适中,该比例直接影响靠近边界的电荷到边界的距离,大量实验表明, Q_2 取 0.1 且 Q_1 取 1 比较合理,而且方便计算。

参 考 文 献

- 靳国栋,刘衍聪. 距离加权反比插值法和克里金插值法的比较[J]. 长春工业大学学报:自然科学版,2003,24(3):53~57.
Jin Guodong, Liu Yancong. Comparison between inverse distance weighting method and Kriging [J]. Journal of Changchun University of Technology: Natural Science Edition, 2003, 24(3):53~57. (in Chinese)
- Stevens Jr D L. Spatial properties of design-based versus model-based approaches to environmental sampling[C]//Caetano M, Painho M. 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Lisbon, Portugal; Instituto Geografico Portugues, 2006: 119~125.
- 吴才聪,马成林,张书慧,等. 基于 GIS 的精确农业合理采样与施肥间距研究[J]. 农业机械学报,2004,35(2):80~83.
Wu Caicong, Ma Chenglin, Zhang Shuhui, et al. Research on reasonable distances of soil sampling and fertilizing based on GIS in precision agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2004, 35(2):80~83. (in Chinese)
- Wang Hongbin, Yang Qing, Liu Zhijie, et al. Determining optimal density of grid soil-sampling points using computer simulation[J]. Transactions of the Chinese Society of Agricultural Engineering, 2006, 22(8):145~148.
- 王秀,赵春江,孟志军,等. 精准农业土壤采样栅格划分方法的研究[J]. 土壤学报,2005,42(2):199~205.
Wang Xiu, Zhao Chunjiang, Meng Zhijun, et al. Field soil sampling grids for precision agriculture [J]. Acta Pedologica Sinica, 2005, 42(2):199~205. (in Chinese)
- Yfantis E A, Flatman G T, Behar J V. Efficiency of Kriging estimation for square, triangular, and hexagonal grids[J]. Mathematical Geology, 1987,19(3):183~205.
- Van Groenigen J W, Stein A. Constrained optimization of spatial sampling using continuous simulated annealing [J]. Journal of Environmental Quality, 1998,27(5):1078~1086.
- Van Groenigen J W, Siderius W, Stein A. Constrained optimization of soil sampling for minimization of the Kriging variance [J]. Geoderma, 1999,87(3~4):239~259.
- Simbahan G C, Dobermann A. Sampling optimization based on secondary information and its utilization in soil carbon mapping [J]. Geoderma, 2006,133(3~4):345~362.
- 张亮,王凌,郑大钟. 有限计算量下模拟退火算法的参数序优化[J]. 控制与决策,2004,19(2):226~229.
Zhang Liang, Wang Ling, Zheng Dazhong. Parameter ordinal optimization for simulated annealing with limited computational efforts[J]. Control and Decision, 2004, 19(2):226~229. (in Chinese)
- 姜成晟,王劲峰,曹志冬. 地理空间抽样理论研究综述[J]. 地理学报,2009,64(3):368~380.
Jiang Chengsheng, Wang Jingfeng, Cao Zhidong. A review of geo-spatial sampling theory [J]. Acta Geographica Sinica, 2009, 64(3):368~380. (in Chinese)

(上接第 123 页)

- 张春梅,刘荣厚,易维明,等. 玉米秸秆等离子体热裂解液化实验[J]. 农业机械学报,2009,40(8):96~99.
Zhang Chunmei, Liu Ronghou, Yi Weiming, et al. Experiment on plasma pyrolysis of corn stalk for liquid fuel [J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(8):96~99. (in Chinese)
- 李大中,张瑞祥,韩璞. 生物质气化过程神经网络模型拟合方法[J]. 太阳能学报,2008,29(5):539~543.
Li Dazhong, Zhang Ruixiang, Han Pu. The method on simulation of biomass gasification with a neural network model[J]. Acta Energetica Solaris Sinica, 2008, 29(5):539~543. (in Chinese)
- 张春梅,易维明,刘庆玉,等. 流化床等离子体加热稳定性试验[J]. 农业机械学报,2010,41(增刊):121~123.
Zhang Chunmei, Yi Weiming, Liu Qingyu, et al. Experiment on the stability of plasma heating fluidized bed [J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(Supp.):121~123. (in Chinese)