

基于DBSCA聚类算法优化的 均值漂移算法的移动通信网络站址 规划优化算法研究

杨锋勇, 宗俭丰, 马 阳

黑龙江科技大学, 电子与信息工程学院, 黑龙江 哈尔滨

收稿日期: 2023年2月8日; 录用日期: 2023年3月4日; 发布日期: 2023年3月13日

摘 要

为适应当今无线通信产业的快速发展, 解决基站选址是否合理, 本文提出了一种基站定位算法, 结合均值漂移算法和DBSCAN聚类算法来解决基站的合理选址问题。该算法将大区域划分为较小的子区域, 并利用均值漂移算法计算每个子区域内局部服务体密度的密度极值点。根据每个高密区域的大小, 建立覆盖范围不同的基站。采用DBSCAN聚类算法对近距离弱覆盖区域进行聚类, 优化传统均值漂移算法的收敛速度。实验结果表明, 该算法在寻找基站定位和数据分类的最优解方面具有实用性和有效性。

关键词

均值漂移算法, DBSCA聚类算法

Research on Optimization Algorithm for Mobile Communication Network Site Planning Based on DBSCA Clustering Algorithm and Optimized Mean Shift Algorithm

Fengyong Yang, Jianfeng Zong, Yang Ma

School of Electronic and Information Engineering, Heilongjiang University of Science and Technology, Harbin
Heilongjiang

Received: Feb. 8th, 2023; accepted: Mar. 4th, 2023; published: Mar. 13th, 2023

文章引用: 杨锋勇, 宗俭丰, 马阳. 基于DBSCA聚类算法优化的均值漂移算法的移动通信网络站址规划优化算法研究[J]. 应用数学进展, 2023, 12(3): 847-859. DOI: 10.12677/aam.2023.123087

Abstract

To adapt to the rapid development of the wireless communication industry and solve the problem of whether a base station location is reasonable, this paper proposes a base station positioning algorithm that combines the mean shift algorithm and the DBSCAN clustering algorithm to address the issue of rational base station selection. This algorithm divides a large area into smaller sub-regions and uses the mean shift algorithm to calculate the density extreme points of the local service body density in each sub-region. Based on the size of each high-density area, base stations with different coverage ranges are established. The DBSCAN clustering algorithm is used to cluster the weak coverage areas at close range, optimizing the convergence speed of the traditional mean shift algorithm. Experimental results show that this algorithm is practical and effective in finding the optimal solutions for base station positioning and data classification.

Keywords

Mean Shift Algorithm, DBSCAN Clustering Algorithm

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

基站选址是通信网络规划的一项重要内容, 在综合考虑信道质量、建设成本、覆盖约束以及其他约束条件下, 规划基站数量和位置[1]。在实际网络规划中, 考虑基站的建设成本和一些其他因素, 有时候可能无法把所有弱覆盖区域都解决, 这时候就需要考虑通信业务量的因素, 尽量优先解决通信业务量高的弱覆盖区域。对于大范围区域基站的选址, 一般建立两种基站, 宏基站和微基站, 选址要达到区域业务量为总业务量的百分之九十, 基站建设成本最低即为合理选址。

对基站进行选择, 两种基站的选择只有选择和不选择, 因此可建立 0~1 整数规划模型。基站选址中优先选择服务量最大的基站, 即在大量数据中寻找聚簇点, 均值漂移算法适用于对低维数据求解目标函数。但是考虑到基站选址过程中智能算法往往是随机生成初始解集[2], 在实际场景中, 通信业务量分布是不均匀的, 比如校园区域的业务量往往较大, 在基站的选址中要优先考虑此类区域。传统的均值漂移算法采用随机生成初始解集的方法在面对大量的不均匀数据集的时候会导致收敛速度变慢, 甚至陷入局部最优[3] [4]。

为解决上述问题, 本文提出一种基于聚类算法优化的均值漂移算法的移动通信网络站址规划优化算法。引入聚类策略, 对信号弱覆盖点进行区域聚类, 把距离近的弱覆盖点聚成一类, 可以得到弱覆盖区域, 对不同的弱覆盖区域分开管理使得可以更好的解决弱覆盖问题。试验结果表明: 该方法可以以较低的建设成本满足基站选址的覆盖条件, 且均值漂移算法的收敛速度更快, 对基站的站址规划有着普适性。

2. 基本概念

为了方便理解基站选址原理和本文提出的算法, 本节介绍以下原理和概念。分别是基站选址基本要求、均值漂移算法、DBSCAN 聚类算法。

2.1. 基站选址原理

对基站进行选择, 基站的选择只有选择和不选择, 属于 0-1 背包规划问题, 在一定的重量范围内选择价值最高的物品, 即优先选择服务量最大的基站, 并通过启发式算法求解目标函数。为了便于计算, 将通信业务量区域用很小的栅格进行划分, 只考虑每个栅格的中心点, 即任给一个区域, 都可以划分成有限个点。每个点有一些属性值, 包括: 坐标, 是否为弱覆盖点, 业务量等。站址也只能选择区域内的点。设选择基站的覆盖范围为 d , 基站所规划的点的坐标为: $P_0(x_0, y_0)$, 则对于坐标为 $P(x, y)$ 的点, 若 $\|P - P_0\| \geq d$, 则认为该点被该基站覆盖, 否则认为该点没有被该基站覆盖。同时, 实际中还需要考虑一个约束条件, 即新建站址之间以及新建站址和现有站址之间的距离不能小于等于基站之间的门限。

2.2. 均值漂移算法

均值漂移算法可以用于聚类, 因为每个均衡都有其自己有吸引盆地。设想在一个有 N 个样本点的特征空间初始确定一个中心点, 计算在设置的半径为 D 的圆形空间内所有的点与中心点的向量计算整个圆形空间内所有向量的平均值, 得到一个偏移均值将中心点移动到偏移均值位置重复移动, 直到满足一定条件结束[5]。

2.3. DBSCAN 聚类算法

DBSCAN 算法是一种基于密度的聚类算法, 它主要原理是通过统计每个点邻域内包含的点的个数来确定该点的密度, 不像 GMM (Gaussian Mixture Models) 需要对数据进行模型假设, 因而它可以发现任意形状的簇。另外, 此算法也不需要复杂的数学计算, 适用于处理大量的数据, 这种算法可以较好地将弱覆盖点进行聚态分类。若两个弱覆盖点的距离不大于 d , 则这两个弱覆盖点应聚为一类, 并且考虑聚类性质具有传递性, 即若点 A 和点 B 是一类的, 点 B 和点 C 是一类的, 则点 A 、 B 和 C 都是一类的。

3. 基站选址模型建立

基站网络覆盖模型的优化描述: $\langle X, F \rangle$ 是由有限集 X 及 F 的一个子集族 F 组成, X 为基站覆盖区域集合, F 为移动通信网络基站覆盖的区域集合[6]。子集族 F 的基站范围覆盖了有限集中的区 X 域, X 中每一个服务区域至少属于 F 中的一个子集。即弱覆盖点 $\subseteq X$, 基站点 $\subseteq F$ 。如何确定移动网络基站的选址位置, 就是找出移动通信网络基站覆盖区域 F 中区域 X 的最小子集 C^* , 使得

$$|C^*| = \min \{ |C| \mid C \subseteq F \text{ 且 } C \text{ 覆盖了 } X \}。$$

假设共选择 n 个位置新建宏基站和微基站, 宏基站覆盖半径 30, 微基站覆盖半径 10, w_j 为基站 j 的服务量, 最大的总服务量为 W , 须达到的总业务量为 Sum 。引入 0-1 变量: p_i, q_i 可得

$$p_i = \begin{cases} 1, & \text{在点 } i \text{ 处建立宏基站} \\ 0 & \text{否则} \end{cases} \quad (1)$$

$$q_i = \begin{cases} 1, & \text{在点 } i \text{ 处建立微基站} \\ 0 & \text{否则} \end{cases} \quad (2)$$

同理引入 0-1 变量, 表示处 j 的弱覆盖点是否被建立在 i 处的新基站覆盖, 可得

$$c(i, j) = \begin{cases} 1, & \text{在 } j \text{ 处的弱覆盖点可以被 } i \text{ 基站覆盖} \\ 0 & \text{否则} \end{cases} \quad (3)$$

其中 p_i 代表宏基站, q_i 代表微基站。

设 $d(i, j)$ 为点 i 和点 j 之间的欧式距离, 可得其表达式为

$$d(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{4}$$

令 $v(i)$ 表示第 i 个基站的业务量, 令 $V(i, j)$ 表示前 i 个基站中能选择的业务量为 j 的最大的总业务量。

如果选择第 i 个基站的业务量大于所需要的总业务量时, 及 $j < w_i$ 时, 不选择 i 基站, 此时选择前 i 个基站的总业务量和选择前 $i-1$ 个基站的总业务量是一样的, 即

$$V(i, j) = V(i-1, j) \tag{5}$$

当选择第 i 个基站的业务量小于所需要的总业务量时, 选择第 i 个基站时的业务量与前 $i-1$ 个基站的总业务量并没有达到所期望的最优总业务量, 所以在选择第 i 个基站和不选择第 i 个基站之间的总服务量之间选取最优结果, 即

$$V(i, j) = \max\{V(i-1, j), V(i-1, j-w_i) + v(i)\} \tag{6}$$

其中的 $V(i-1, j)$ 表示不选择第 i 个基站的总业务量, 即选择前 $i-1$ 个基站的业务类。 $V(i-1, j-w_i) + v(i)$ 表示选择了第 i 个基站, 能选择的最大业务量减少了 w_i , 但是总业务量增加了 $v(i)$ 。

由此可得递推关系式

$$V(i, j) = V(i-1, j) \quad j < w_i \quad (w_i \text{ 为基站 } i \text{ 需要提供的服务量}) \tag{7}$$

$$V(i, j) = \max\{V(i-1, j), V(i-1, j-w_i) + v(i)\} \quad j \geq w_i \tag{8}$$

可以得到初始条件 $V(i, 0) = V(0, j) = 0$, 即选择前 i 个基站但服务容量为 0 和选择 0 个基站但业务量为 j 的背包价值都为 0; 确定好动态规划函数后, 考虑函数每一次迭代的情况。

当函数迭代完成之后, $V(n, W)$ 为总业务量为 W 的背包选择前 n 个基站。

通过上面的模型可以得到基站建设成本的最优解, 还需要知道选择了哪些基站, 所以进行回溯找到解的组成。

$V(i, j) = V(i-1, j)$ 时, 说明没有选择第 i 个基站, 则回到 $V(i-1, j)$; $V(i, j) = V(i-1, j-w_i) + v(i)$ 时, 说明选择了第 i 个基站, 该基站为最优解组成的一部分, 随后回到没有选择该基站之前, 即回到 $V(i-1, j-w_i)$; 一直遍历到 $i=0$ 结束为止, 所有基站的组成都可以解得出来。

实际情况中, 每个基站的信号覆盖并不是完全的圆形覆盖, 而是每个站上有三个扇区, 每个扇区指向一个方向。每个扇区在主方向上覆盖范围最大(半径方向), 在主方向左右六十度的范围内可以覆盖, 覆盖范围按线性逐渐缩小, 在六十度的时候, 覆盖范围为主方向覆盖范围的一半。超过六十度, 则无法被该扇区覆盖。需要考虑每个站的任意两个个扇区的主方向之间的夹角不能小于四十五度, 考虑现网基站的坐标点, 新建站址之以及新建站址和现有站址之间的距离的门限是 10。

设主方向的方向轴顶点上的坐标为 (x_k, y_k) , 已知基站坐标为 (x_i, y_i) , 根据欧式距离公式可得微基站扇形主方向最高点欧式距离为 s_1 , 宏基站扇形主方向最高点欧式距离为 s_2 , 则

$$(x_i - x_k)^2 + (y_i - y_k)^2 = 10^2 \quad \text{微基站扇形主方向顶点距离} \tag{9}$$

$$(x_i - x_k)^2 + (y_i - y_k)^2 = 30^2 \quad \text{宏基站扇形主方向顶点距离} \tag{10}$$

设弱覆盖点 (x_j, y_j) 与原点的欧式距离为 s_3 , 弱覆盖点与基站扇形主方向最高点的欧式距离为 s_4 . 可得

$$s_3 = \sqrt{x_j^2 + y_j^2} \tag{11}$$

$$s_4 = \sqrt{(x_j - x_k)^2 + (y_j - y_k)^2} \tag{12}$$

设 s_3 和 s_1 之间的夹角为 θ ，由余弦公式和反三角公式可得

$$\theta = \arccos\left(\frac{s_1^2 + s_3^2 - s_4^2}{2 \cdot s_1 \cdot s_3}\right) \quad (13)$$

每个站的任意两个扇区的主方向之间的夹角不能小于 45° ，设每个站的任意两个扇区的主方向之间的夹角为 α ，假设第 i 个扇形的主方向直线斜率为 k_i ，第 j 个扇形的主方向直线斜率为 k_j ，可得

$$k_i = (y_k - y_i)/(x_k - x_i) \quad (14)$$

$$k_j = (y_k - y_j)/(x_k - x_j) \quad (15)$$

两条直线之间的角度即两条直线之间的转向角，由正弦公式和反三角公式可得

$$\alpha = \arctan \frac{k_2 - k_1}{1 + k_1 \cdot k_2} \quad (16)$$

由已知在主方向左右 60° 的范围内可以覆盖，覆盖范围按线性逐渐缩小条件，且在 60° 的时候，覆盖范围为主方向覆盖范围的一半这两个条件，我们可以得到与主方向轴尖角 60° 度角半径 R 的线性表达式，设

$$R = -k\theta + r \quad \theta \subseteq [0, 2\pi] \quad (17)$$

其中 k 为半径 R 的线性衰减系数， θ 为半径 R 的自变量， r 为基站扇区主方向最大覆盖距离，宏基站时 $r = 30$ ，微基站时 $r = 10$ 。由在 60° 时覆盖范围为主方向覆盖范围的一半该条件我们可以知道当该基站为宏基站时的 $\theta = \pi/3$ ， $r = 30$ ，解得

$$R_{\text{宏}} = -\frac{45}{\pi} \cdot \theta + 30 \quad \text{宏基站半径衰减公式} \quad (18)$$

$$R_{\text{微}} = -\frac{15}{\pi} \cdot \theta + 10 \quad \text{微基站半径衰减公式} \quad (19)$$

进一步考虑，把扇区的覆盖形状当成三棱形处理，以基站的站址为旋转中心进行三棱形旋转，旋转过程中以扇区覆盖的总业务量最大作为扇区旋转的停止点。设扇区旋转角度为 φ ，则半径衰减公式变化为

$$R = -k \cdot (\theta - \varphi) + r \quad \varphi \subseteq \left(-\frac{\pi}{3}, \frac{\pi}{3}\right) \quad (20)$$

在基站上任意取一个扇区做旋转变换，假设扇区主方向旋转角为 φ 时得到第一个扇区的业务量最大，即得到第一个扇区位置， $\varphi \subseteq \left[0, \frac{2\pi}{3}\right]$ 。第二个扇区的 φ_1 取值变化为 $\varphi_1 \subseteq \left[\frac{\pi}{4} + \varphi, \frac{11\pi}{12} + \varphi\right]$ ；同理，当第二个扇区在旋转过程中得到覆盖区域内最大的业务量停止旋转，即确立了第二个扇区的位置；第三个扇区的 φ_2 取值变化为 $\varphi_2 \subseteq \left[\frac{\pi}{2} + \varphi_1, \frac{7\pi}{6} + \varphi_1\right]$ 。扇区经过三次的旋转求和对比，便可以确立下每个基站的三个扇区最佳位置。

综上所述，我们可以得到基站选址基本模型：

$$\text{MAX} \quad W = \sum_{i=0, j=0}^n (w_i p_i + w_j q_i) \quad (21)$$

$$\text{MIN} \quad Y = \sum_{i=0, j=0}^n (10p_i + q_i) \quad (22)$$

s.t

$$\sum_{i=0}^n c(i, j) \cdot p_i \geq 1, \sum_{i=0}^n c(i, j), 0 \leq j \leq n \tag{23}$$

$$c(i, j) \cdot (R - d(i, j)) \geq 0, 0 \leq i \leq n, 0 \leq j \leq m \tag{24}$$

$$(1 - c(i, j)) \cdot (d(i, j) - R) \geq 0, 0 \leq i \leq n, 0 \leq j \leq m \tag{25}$$

$$p_i = 0, 1 \tag{26}$$

$$q_i = 0, 1 \tag{27}$$

$$\sum_{i=1, j=1}^n w_j \cdot p_i + \sum_{i=1, j=1}^n w_j \cdot q_i \geq Sum \tag{28}$$

$$\sqrt{(x_i - x_j)^2 - (y_i - y_j)^2} \leq 30, (x_i, y_i \text{ 属于宏基站}) \tag{29}$$

$$\sqrt{(x_i - x_j)^2 - (y_i - y_j)^2} \leq 10, (x_i, y_i \text{ 属于微基站}) \tag{30}$$

$$\sqrt{(x_{i+1} - x_i)^2 - (y_{i+1} - y_i)^2} \leq 10 \tag{31}$$

$$\theta \leq 60 \tag{32}$$

$$\alpha \geq 45 \tag{33}$$

4. 基于 DBSCA 聚类算法优化的均值漂移算法

4.1. DBSCAN 算法分析

DBSCAN 算法是一种基于密度的聚类方法，可以找到样本点的全部密集区域，并把这些密集区域当作一个的聚类簇。如图 1 所示，相同颜色代表是在同一个聚类簇。DBSCAN 的优势之一是它能够找到任意形状的聚类，这与其他聚类算法(例如 k-means 聚类算法)不同，后者假设聚类是球形的。此外，DBSCAN 不需要事先指定聚类的数量，使其成为探索性数据分析的合适算法。

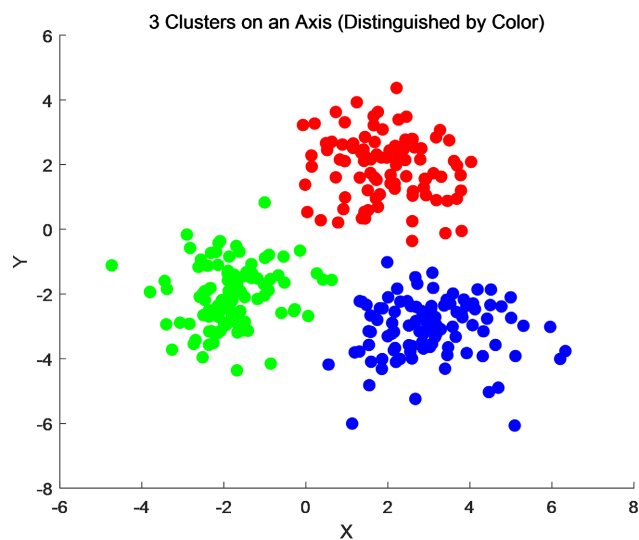


Figure 1. Each dense area clustered cluster graph

图 1. 每个密集区域聚类簇图

领域半径和最少点数目这两个参数可以刻画密集程度，当领域半径 R 内的点的个数大于最少点数目时，就是密集。如图 2 中圆形区域即为领域半径 $R=1$ ，最少点数目为 13。

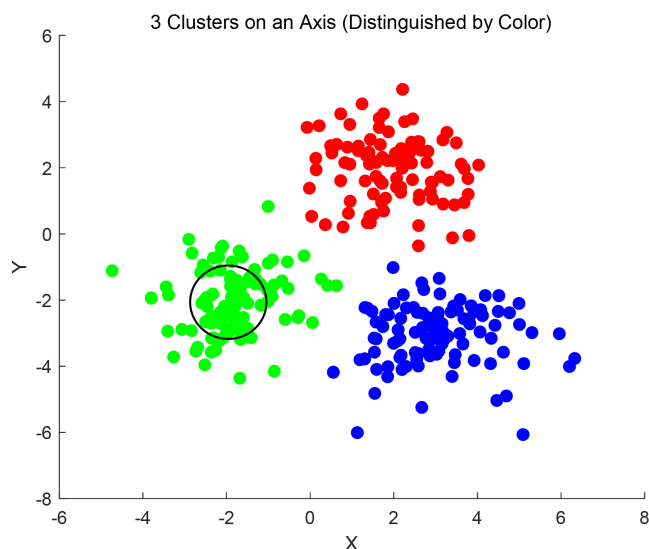


Figure 2. Two algorithm parameters diagram in DBSCAN
图 2. DBSCAN 中 2 个算法参数图

样本点分为核心点，边界点和噪声点。领域半径 R 内的样本点数量大于等于最少点目的点称为核心的点，不属于核心点但在某个核心点的领域内的点属于边界点，既不是核心点也不是边界点的样本点为噪声点。如果 P 为核心样本点， Q 在 P 的 R 邻域内，那么称 P 到 Q 密度直达。任何核心点到其自身密度直达，密度直达不具有对称性，如果 P 到 Q 密度直达，那么 Q 到 P 不一定密度直达。如果存在核心样本点 P_1, P_2, \dots, P_n ，且 P_1 到 P_2 密度直达， P_2 到 P_3 密度直达， \dots ， $P_{(n-1)}$ 到 P_n 密度直达， P_n 到 Q 密度直达，则 P_1 到 Q 密度可达。密度可达也不具有对称性。如果存在核心样本点 S ，使得 S 到 P 和 Q 都密度可达，则 P 和 Q 密度相连。密度相连具有对称性，如果 P 和 Q 密度相连，那么 Q 和 P 也一定密度相连。密度相连的两个点属于同一个聚类簇。

如果两个样本点不属于密度相连关系，则两个点非密度相连。非密度相连的两个样本点属于不同的聚类簇，或者其中存在噪声点。

4.2. DBSCAN 算法伪代码

DBSCAN 算法伪代码见表 1。

Table 1. Pseudocode of DBSCAN algorithm
表 1. DBSCAN 算法伪代码

算法 1: DBSCAN 算法

输入: 样本集 $B = \{y_1, y_2, \dots, y_k\}$

领域参数 $(\varepsilon, MinPts)$

过程:

1: 初始化核心对象: $\Psi = \emptyset$

Continued

```

2:   for  $i = 1, 2, \dots, k$  do
3:       确定样本  $y_i$  的  $\varepsilon$ -领域  $N_\varepsilon(y_i)$ 
4:       if  $|N_\varepsilon(y_i)| \geq MinPts$  then
5:           将样本  $y_i$  加入核心对象集合:  $\Psi = \Psi \cup \{y_i\}$ 
6:       end if
7:   end for
8:   初始化聚类簇数:  $z = 0$ 
9:   初始化未访问样本集合:  $K = B$ 
10:  while  $\Psi \neq \emptyset$  do
11:      记录当前未访问样本集合:  $K_{old} = K$ 
12:      随机选取一个核心对象  $g \in \Psi$ , 初始化队列  $R = \langle g \rangle$ 
13:       $K = K \setminus \{g\}$ 
14:      while  $R \neq \emptyset$  do
15:          取出队列  $R$  中的首个样本  $r$ 
16:          if  $|N_\varepsilon(r)| \geq MinPts$  then
17:              令  $X = N_\varepsilon(r) \cap K$ 
18:              将  $X$  中的样本加入队列  $R$ 
19:               $K = K \setminus X$ 
20:          end if
21:      end while
22:       $n = n + 1$ , 生成聚类簇  $S_n = K_{old} \setminus K$ 
23:       $\Psi = \Psi \setminus S_n$ 
24:  end while
输出: 簇划分  $S = \{S_1, S_2, \dots, S_z\}$ 

```

4.3. 均值漂移算法**4.3.1. 算法基础公式**

均值漂移的过程为不断重复计数距离均值, 移动中心点, 设 Sh 为以 x 为中心点, 半径为 h 的二维圆形区域, k 为包含在 Sh 范围内点的个数, x_i 为包含在 Sh 范围内的点, 则偏移均值公式为

$$M(x) = \frac{1}{k} \sum_{x_i \in Sh} (x - x_i) \quad (34)$$

设 M^t 为 t 时间下求得的偏移均值, x^t 为 t 时间下的中心, 则一定距离之后的中心点所处位置计算公式为

$$x^{t+1} = M^t + x^t \quad (35)$$

4.3.2. 算法原理及适用性分析

通过对随机选择的点周围指定半径 r 内的数据进行求和和记录, 从而找到随机局部区域的密度极值

(峰值)的方法。通过上下左右移动指定的距离 ζ 并取五个点之间的最大值重复该过程，直到中心点为最大值。将得到的密度最大点作为基站点，计算覆盖业务量。重复上述过程，直至中心点为最大值，此时该点成为该局部区域的密度极大值点。

在特征空间中初始化随机起点 x_0 ，通过对点周围指定半径内的数据求和来计算的密度：

$$p(x_0) = \sum_{x_i \in Sh} K(x_0, x_i)$$

，半径内的所有数据点 x_i 求和，其中 K 是核函数。将起点 x_0 向最大密度梯度的方向

移动： $x_1 = x_0 + \eta \cdot grad(p(x_0))$ ，其中 η 为步长， $grad(p(x_0))$ 为 x_0 处的密度梯度。直到中心点是最大密度或满足停止条件： $x_{k+1} = x_k + \eta \cdot grad(p(x_k))$ ，最后收敛的点 x_k 为密度极值(峰值)，可以作为基站点。

根据上述原理，可以利用均值漂移算法求出随机局部区域的密度极值点，确立该点为基站点，算出覆盖业务量，更新弱覆盖点、已有基站信息；再次求极值点，将所有新建基站覆盖业务量求和；当业务量达到总业务量的 90% 便停止迭代。

4.3.3. 均值算法流程图

利用聚类之后的弱覆盖点作为随机初始解集进行求解，算法流程图如图 3 所示。

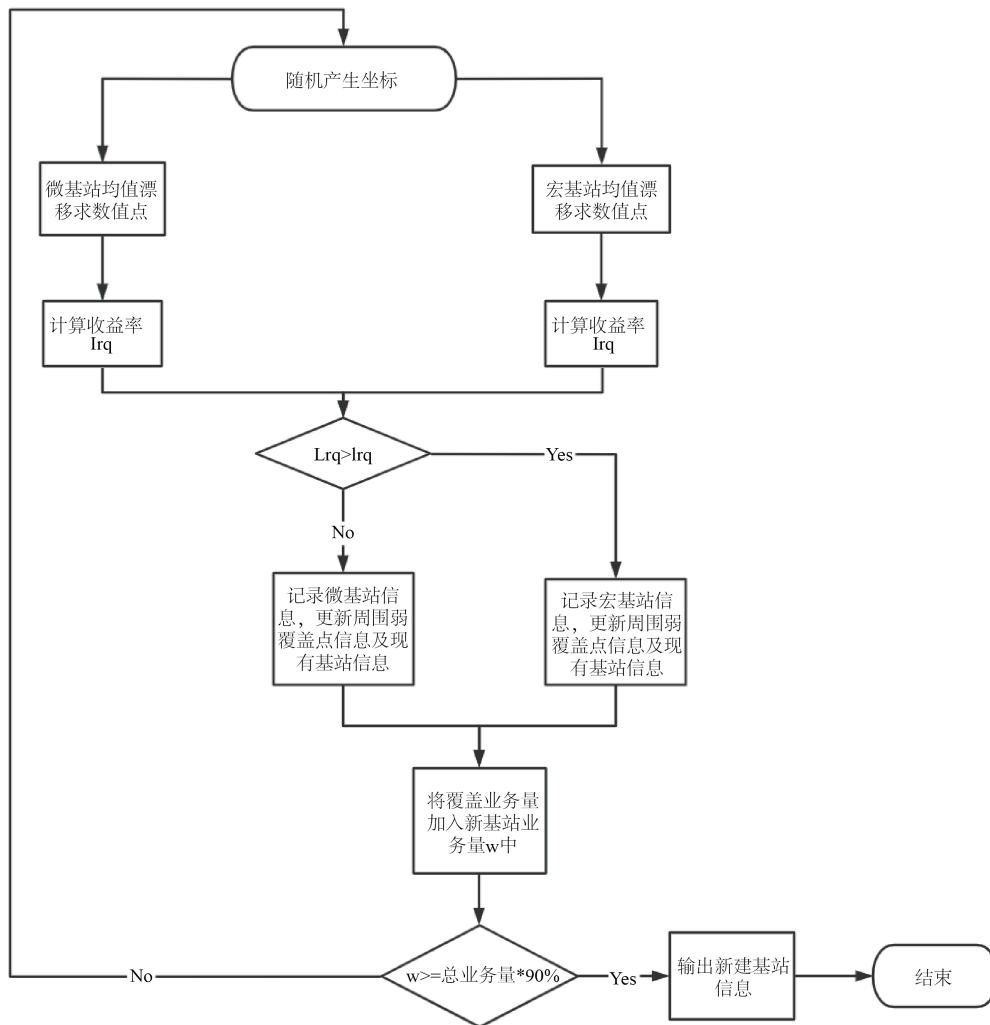


Figure 3. Mean shift algorithm flowchart

图 3. 均值漂移算法流程图

5. 试验分析

5.1. 聚类结果

使用 DBSCAN 算法进行 18 万组弱覆盖点数据的处理之后, 结果用时 1371 秒将数据分成 627 类, 其全体的分类如图 4 表示。图 5~8 表示的为部分聚类簇取样表示, 其间隔为 100。图 9 为某城市某区域的现网覆盖情况, 其中红色的区域表示为弱覆盖区域。

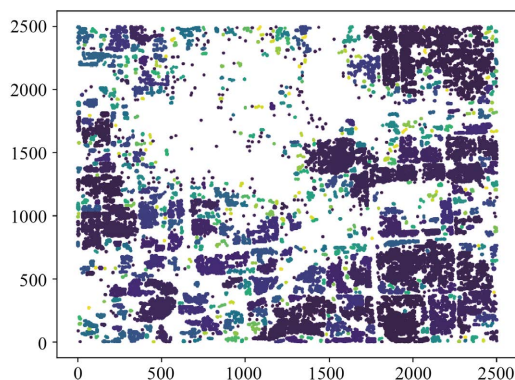


Figure 4. Cluster map of all weak coverage points
图 4. 所有弱覆盖点聚类图

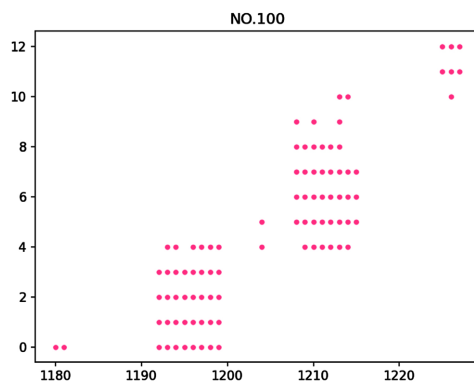


Figure 5. Partial clustering effect image
图 5. 部分聚类效果图

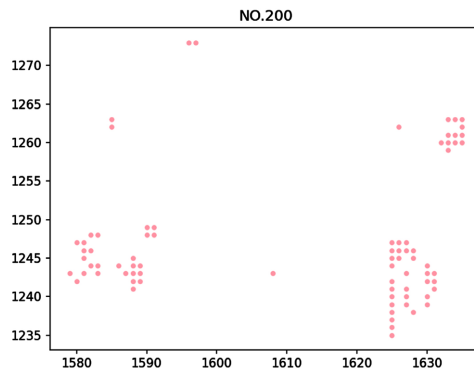


Figure 6. Partial clustering effect image
图 6. 部分聚类效果图

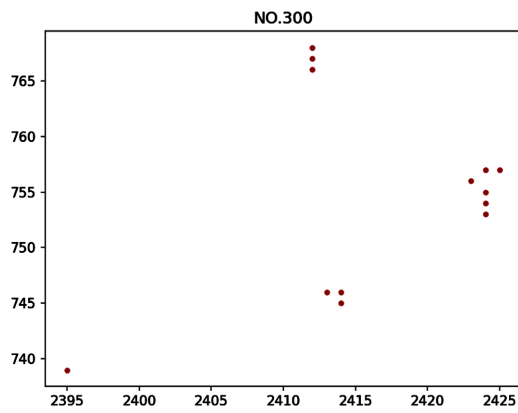


Figure 7. Partial clustering effect image
图 7. 部分聚类效果图

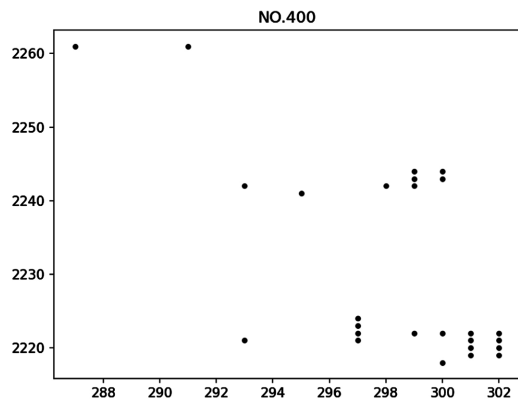


Figure 8. Partial clustering effect image
图 8. 部分聚类效果图

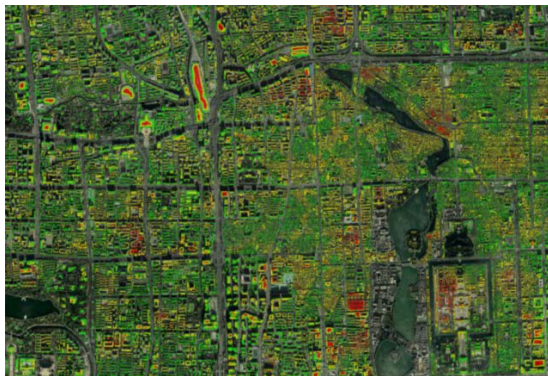


Figure 9. Current network coverage map of a certain area in a city
图 9. 某城市某区域现网覆盖图

使用聚类之后的聚类中心代替原有弱覆盖点，可以得到高质量的均值漂移算法初始解集。

5.2. 均值漂移算法求解模型结果

经过多次使用均值漂移的结果对比，得到宏基站的选址和微基站的选址，其可视化结果如图 10 所示。

图中小圆代表微基站，大圆代表宏基站。

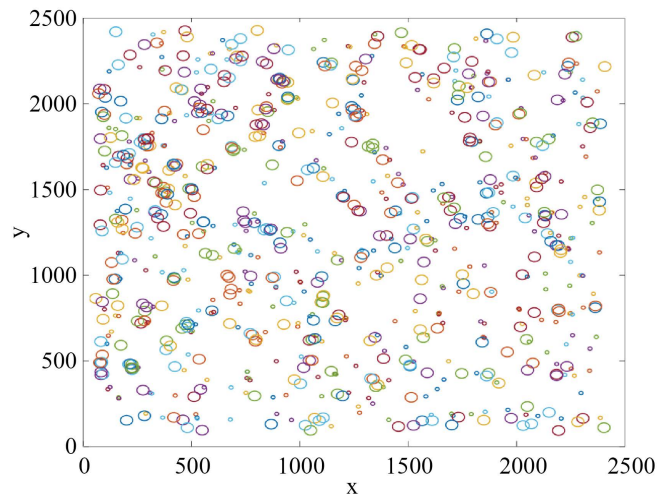


Figure 10. Visualization of base station location selection map
图 10. 基站可视化选址图

5.3. 算法适用性优化

考虑到有两种基站可供选择,确立可在同一随机区域将 r 定义为 10 和 30,对应的 ζ 定义为 20 和 60; 同时进行漂移求极值点, 求出对应收益率 $lr_q = w_j q_j d$ 和 $lr_p = w_j p_j d / 10$ 比较两个收益率, 留下较大的。这样可以尽可能减小成本, 增大效益。

基站选址与已有基站之间的欧式距离必须大于 10,所以在进行均值漂移时,需要判断是否能建基站; 若不能, 则将此点覆盖业务量记为 0。采用 DBSCAN 算法对初始解集进行聚类之后, 均值漂移算法的迭代次数明显下降, 传统均值漂移算法和优化后均值漂移算法效果对比图如图 11 所示。

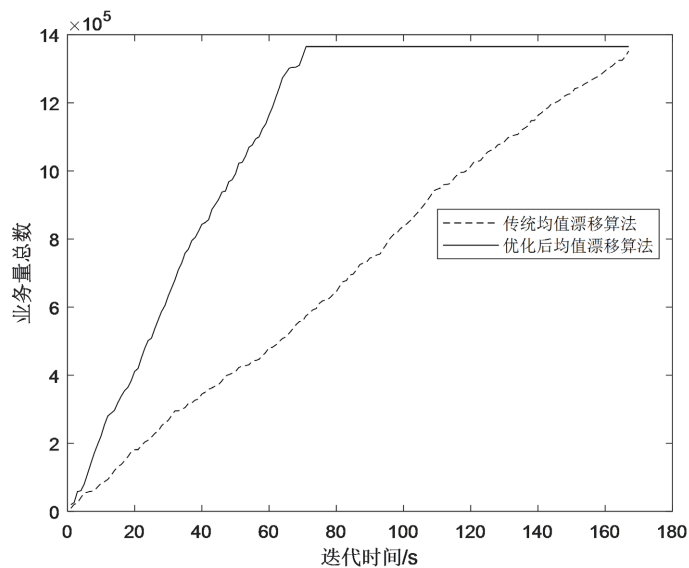


Figure 11. Comparison of the effects of two algorithms image
图 11. 两种算法效果比较图

6. 结论

基站选址是当今通信服务的一个重要环节, 本文对基站选址问题进行了分析, 以区域总业务量达到最大以及基站建设成本最低为指标建立了最优化模型, 并提出一种基于 DBSCAN 聚类算法优化后的均值漂移算法, 求解最优基站建设方案。试验结果表明, 本文提出的方法能有效提高优化效率, 对基站选址具有一定的实用性。

参考文献

- [1] 黄骅, 江俊. 基于聚类的无线网络基站选址优化算法研究[J]. 现代信息科技, 2018, 2(9): 50-52.
- [2] 苏延平. 依赖混合型位置大数据的均值漂移聚类算法[J]. 山西能源学院学报, 2020, 33(2): 97-99.
- [3] 张双寒, 王珊, 王特. 基于 DBSCAN 聚类的毕星团成员星识别方法[J]. 现代信息科技, 2021, 5(24): 146-149.
- [4] 张乾, 张强. 动态规划迭代算法在末端防御中的应用[J]. 电子设计工程, 2021, 29(3): 104-107.
- [5] 赵华茗, 余丽, 周强. 基于均值漂移算法的文本聚类数目优化研究[J]. 数据分析与知识发现, 2019, 3(9): 27-35.
- [6] 陈琦, 贾元华. 基于人工鱼群算法的移动通信网络基站覆盖优化问题[J]. 北京交通大学学报, 2013, 37(6): 99-102.