ISSN: 2664-9640

Research on communication base station location based on DBSCAN clustering algorithm

Shumiao Xie*, Nan Li, Yuhan Sun

Sichuan University-Pittsburgh Institute, Sichuan University, Chengdu, 610207, China

* Correspondence: 2019141520165@stu.scu.edu.cn

Abstract

By analyzing the unit cost of macro base station and micro base station, this paper selects all micro base stations to reduce the cost. Then, the coordinates of the new base station are determined with the objective function of maximizing the traffic and minimizing the total cost. On this basis, all weak coverage points are clustered to form a weak coverage point region. DBSCAN clustering algorithm based on density and low time complexity is adopted. Divide all weak coverage points into different clusters, and finally get the number of clusters and the number of weak coverage points in each cluster.

Keywords

Maximize business volume; Minimum total cost; DBSCAN clustering.

1. Introduction

With the development of science and technology, mobile communication is becoming more and more popular, and communication technology is also an important embodiment of national strength. Since 5G technology was put into market application, the bandwidth of communication has become larger and larger, but the coverage of base station has become smaller and smaller. Therefore, for the same area, more communication base stations need to be established to meet the needs of more people [1].

This paper analyzes the network coverage of a certain area and analyzes the location. Under different circumstances, different location schemes will be proposed. By establishing base stations at weak coverage points, we can achieve the goal of high coverage, maximize the traffic and save the cost.

2. Model establishment

2.1. Coordinates of the new base station

Firstly, the first cycle is established to calculate the distance between the positions of all weak coverage points and the positions of existing base stations:

$$d_i = \sqrt{(x_i - x_{built})^2 + (y_i - y_{built})^2} \leq 10 \tag{1}$$

Where d_i represents the distance between the ith weak coverage point and the established base station, x_i and y_i represents its coordinates. x_{built} and y_{built} respectively indicates the coordinates of the established base stations.

$$d_i = \sqrt{(x_i - x_{new})^2 + (y_i - y_{new})^2} \leq 10 \tag{2}$$

$$d_{i\_cannot} = \sqrt{(x_{i\_cannot} - x_{new})^2 + (y_{i\_cannot} - y_{new})^2} \leq 10 \tag{3}$$

$$traffic_{covered} = traffic_{new} + traffic_i + traffic_{i\_cannot} \tag{4}$$

Where x_{new} and y_{new} indicates the coordinates of the new building base station(BTS) coordinates, x_{cannot} , y_{cannot} cannot respectively indicates that the location coordinates of the new base station cannot be established, $d_{i_{cannot}}$ indicates the distance between the location where the new base station cannot be established and the new base station, $traffic_{covered}$ refers to the traffic covered after the new base station is built. $traffic_i$ represents the traffic volume of the ith weak coverage point, $traffic_{i\_cannot}$ indicates the traffic volume of the location that cannot establish a new base station but is covered by the new base station. Then remove all overwritten locations from the original list. Then return to the second cycle and judge:

$$traffic_{covered} \geq 90\% traffic_{all} \tag{5}$$

Where $traffic_{all}$ indicates the total amount of business at all initial weak coverage points. If the condition is not satisfied, the loop continues. If the condition is satisfied, the loop jumps out. Therefore, the final cycle will stop when the new base station is just 90%, which can also reduce the cost.

A total of 4124 new base stations need to be built, all of which are micro base stations. The total amount of coverage services after the establishment of new base stations is 6350743.136, accounting for 90.002% of the total traffic of weak coverage points. The cost of establishing new base stations is 41240.

The coordinates of the new BTS are shown in Table 1：

Table 1 New BTS site coordinates

x coordinate	y coordinate	x coordinate	y coordinate
1356	2271	1019	1593
869	2292	935	1665
881	1256	1489	1228
1096	1658	699	2014
683	2198	972	1238
865	2012	1314	1125
1335	2206	753	1297
1357	1086	892	2001
675	1957	1229	1894
932	2210	1000	1858

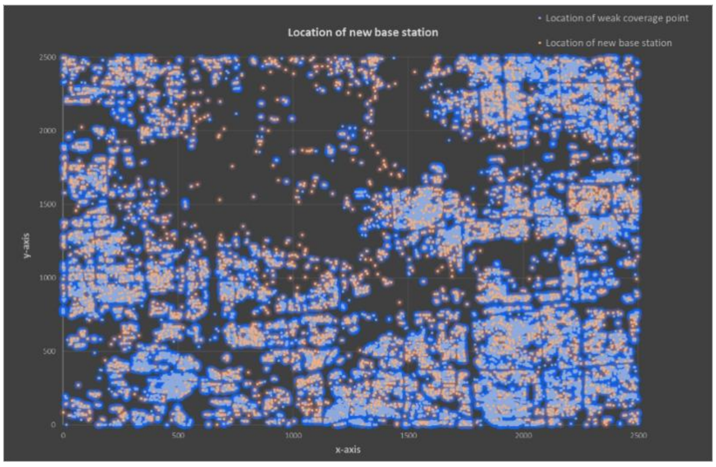

Figure 1 Location diagram of new base station

2.2. Calculate the proportion of business volume covered

Based on the optimization results in Section 2.1, this problem discusses the optimal coverage obtained by determining the angles of the three main directions of the new station when the coverage effect of the base station cannot reach the ideal situation, that is, when the left and right coverage radius of the three main directions of each station decreases linearly.

Establish a plane rectangular coordinate system with the base station as the coordinate origin (x_0, y_0). θ is the angle between a point (x_1, y_1) and the positive direction of the coordinate y axis

$$\theta = \arctan\left(\frac{x_1 - x_0}{y_1 - y_0}\right), \qquad \theta \in [0,2\pi] \tag{6}$$

For a base station, a certain main direction $\vec{N} = (x_n, y_n)$, any angle within the maximum coverage angle θ_s. The ideal coverage radius is R, and the distance from the farthest coverage point to the base station, that is, the coverage radius is L, then:

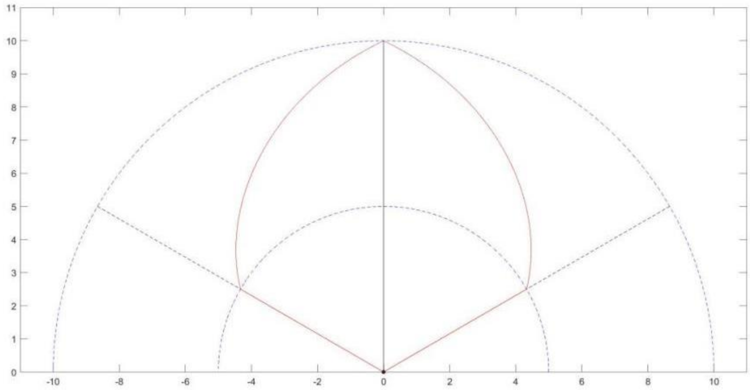

Figure 2 Geometric diagram of single main direction coverage

$$L(\theta_s) = R - \left|\frac{3R}{2\pi} * \theta_s\right|, \qquad \theta \in [-\frac{\pi}{3}, \frac{\pi}{3}] \tag{7}$$

If the three main directions are evenly distributed, the general formula of coverage radius is:

$$L(\theta) = \left|\frac{3R}{2\pi}(\theta - \frac{\pi}{3} - \frac{2\pi n}{3})\right| + \frac{1}{2}R, \quad n = \left\lfloor\frac{3\theta}{2\pi}\right\rfloor, \quad \theta \in [0,2\pi] \tag{8}$$

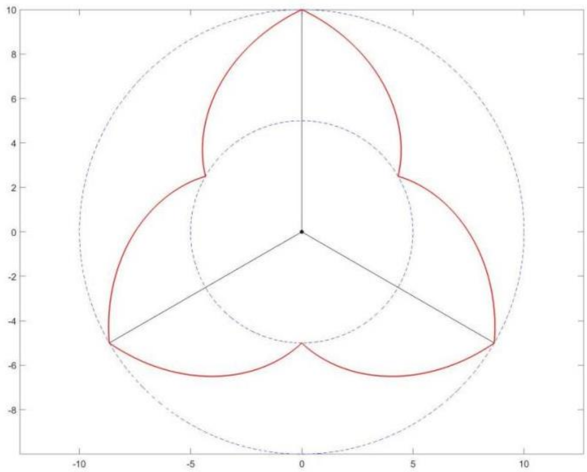

Figure 3 Schematic diagram of coverage in main direction of uniform distribution

Angle between a point (x_1, y_1) and the main direction $\vec{N}$ of the base station θ_d:

$$\theta_d = \left|arctan\left(\frac{x_n - x_0}{y_n - y_0}\right) - arctan\left(\frac{x_1 - x_0}{y_1 - y_0}\right)\right|, \qquad \theta_d \in [0,2\pi] \tag{9}$$

Distance between point (x_1, y_1) and point (x_2, y_2):

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{10}$$

The main directions of 4124 new stations determined by the above algorithm, under the condition of sector coverage, the total traffic that can be covered is 6.0474 * 106, accounting for 85.7% of the total traffic, and the loss is about 4.3% compared with the circular coverage.

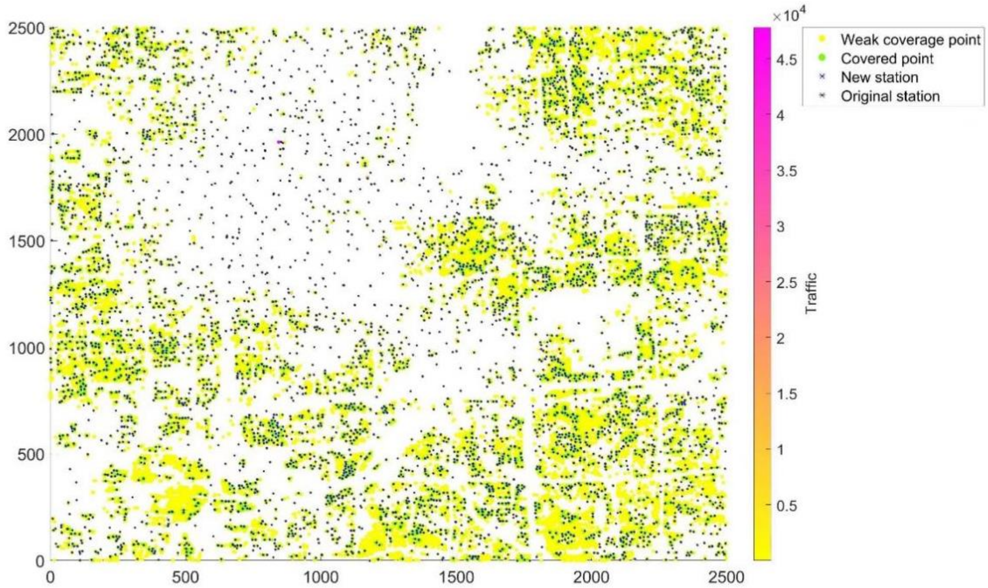

Figure 3 Schematic diagram of coverage

Table 2 New BTS main direction

Site Abscissa	Site Ordinate	Main direction Angle with positive direction of X axis /°	Site Abscissa	Site Ordinate	Main direction Angle with positive direction of X axis /°
0	948	0.00	1	1268	-179.95
0	948	90.00	1	1268	90.00
0	948	161.57	1	1268	35.54
0	1221	180.00	1	1814	-179.97
0	1221	180.00	1	1814	0.00
0	1221	90.00	1	1814	90.00
0	1232	180.00	3	895	-179.81
0	1232	135.00	3	895	180.00
0	1232	80.54	3	895	75.96
0	1673	180.00	4	1279	-45.00
0	1673	0.00	4	1279	180.00
0	1673	53.13	4	1279	56.31
0	2496	180.00	4	1875	-23.20
0	2496	0.00	4	1875	-90.00
0	2496	90.00	4	1875	51.34
1	929	-179.94	5	1379	-126.87
1	929	0.00	5	1379	-63.43

1	929	63.43	5	1379	135.00
1	993	-90.00	5	2425	-90.00
1	993	0.00	5	2425	180.00
1	993	90.00	5	2425	108.43
1	1073	-90.00	6	1703	-179.80
1	1073	90.00	6	1703	26.57
1	1073	139.40	6	1703	116.57

2.3. Weak coverage region clustering

Based on the requirements of distance and transitivity, this paper selects DBSCAN algorithm to cluster weak coverage points [2,3]. Firstly, the weak coverage grid data is imported, and the abscissa and ordinate of each point are represented by X and Y respectively.

Next, we need to cluster the black points (weak coverage points) in the figure above. The points with distance of less than 20 are clustered into one class, and the transitivity is considered. It should be noted that since there is no difference between weak coverage points, this paper believes that if there is a weak coverage point, it can be grouped into one class, that is, the influence of noise is not considered.

Based on the powerful statistical and machine learning toolbox of MATLAB, this paper calls DBSCAN function to get the final clustering result [4,5].

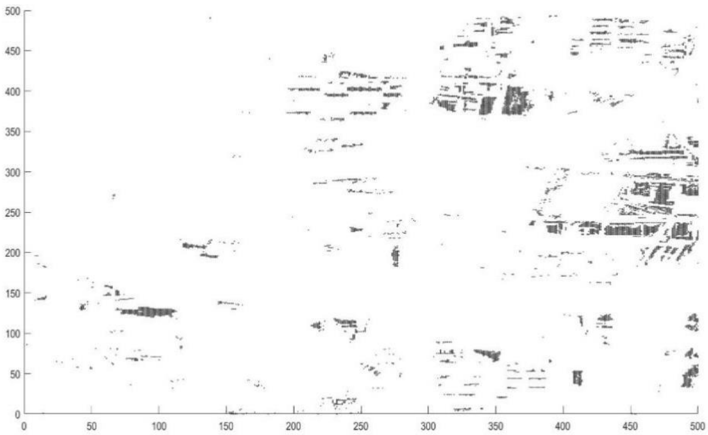

Figure 4 Weak covering scatter

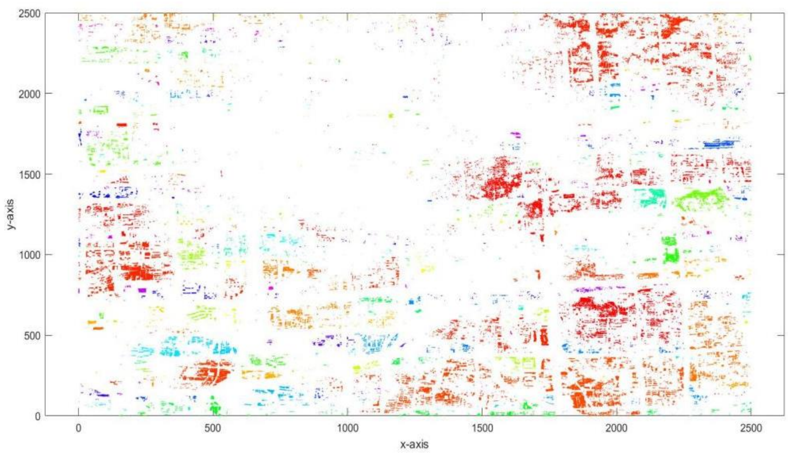

Figure 5 Schematic diagram of clustering results

3. Model evaluation and improvement

3.1. Merit model

1. Simplify complex problems. By analyzing the problem, it is transformed into a simple problem, and then modeling and programming, so as to complete the task with less operation and reduce redundancy.

2. Strong operability. As long as the location coordinates and traffic volume of all points are obtained, the results can be obtained without additional data or conditions, and the operation is simple.

3.2. Model shortcomings

The use of this model is limited. Since this model is based on the topic, if the unit cost of macro base station is lower than that of micro base station, the model needs to be modified before use.

References

[1] Keyu Li, National Development and Reform Commission: 81000 5g base stations were built in the first quarter of this year to ensure that the construction of 1 billion mu of high standard farmland in China will be completed on schedule by the end of the year.

[2] Xuan Ma, Anomaly detection method of viewing data set based on improved DBSCAN algorithm and K-S test [D] Beijing University of Posts and telecommunications, 2021.

[3] Lulu Li, Point cloud segmentation based on adaptive DBSCAN clustering [D] Liaoning University of engineering and technology, 2021.

[4] Xianxin Lin, Chengbao Ye, Stability analysis of sand core holder system based on MATLAB [J] Times automobile, 2022 (06): 150-151.

[5] Hao Ning, Yan Zhu, Jian Zhou, Internal model control internal combustion engine fault analysis system based on MATLAB [J] Metalworking (cold working), 2022 (03): 84-88.