ISSN: 2710-0723

Using Cluster Analysis to Solve Mobile Communication Network Site Planning Problems

Qiran Zhang, Mengyao Yang, Xiaoyang Liu and Jiaye Wang
School of Anhui University of Finance and Economics Anhui 233030, China

Abstract

The area planning problem involves optimization model and clustering model, according to the site construction requirements, the base station planning needs to cover 90% of the business volume in the weak coverage area. In the process of site selection, it is required that the cost is as low as possible and the constraints of each condition are satisfied. Therefore, according to different site construction conditions, two corresponding planning schemes and a clustering method are derived.

Keywords

Regional Planning; Base Station Construction; K-Means; Maltab; Data Cleaning.

1. Introduction

1.1. Background of the Problem

In the era of 5G development, mobile communication technology is rapidly developing and the communication bandwidth is getting larger, so the problem of insufficient base station coverage comes into play. To solve this situation, the number and type of base stations need to be increased. Therefore the planning of communication network becomes a major solution. Station site selection needs to give weak coverage areas based on the coverage of antennas in the existing network, and then select suitable points to build new base stations to solve the weak coverage problem of the existing network. In the actual planning, the cost issue needs to be considered, and according to the real situation, it is required to give priority to solve the weak coverage area with high service volume.

1.2. Problem Restatement

Problem 1: For the given information of weak coverage points and existing network base stations, a model is required to derive a reasonable station site plan to set up appropriate macro base stations and micro base stations so that 90% of the weak coverage points are covered by the service volume. Also consider the cost issue and satisfy that the distance between any station sites should not be less than 10.

Problem 2: In continuation of Problem 1, in connection with the actual situation, consider that the station site is not a complete circular coverage, but three sector coverage with adjustable sector direction, but the angle between the main directions of any two sectors is required to be greater than or equal to 45 degrees. It is required to calculate the optimal station site and sector angle selected in such a case, and whether the new station can cover 90% of the total service volume.

Problem 3: Build a model to cluster the weak coverage points in the area to get the weak coverage area. The selected model needs to make the time complexity used for clustering as low as possible.

ISSN: 2710-0723

2. Text Analysis

2.1. Problem One

The first Problem actually asks us to we have to create new base stations for area planning. We can treat each base station as the center point of a circle, the radius of the circle where the macro base station is located is 30 and the radius of the circle where the micro base station is located is 10. According to the requirements of the Problem, we can solve the requirements of this problem one if we can establish coordinates within the circle and eliminate duplicate and invalid base stations so that the final circle has more than 90 percent of the points.[1] It is worth noting that we need the special requirements given by the topic: the consideration of cost minimization and the distance of existing base stations cannot be less than 10. [3]

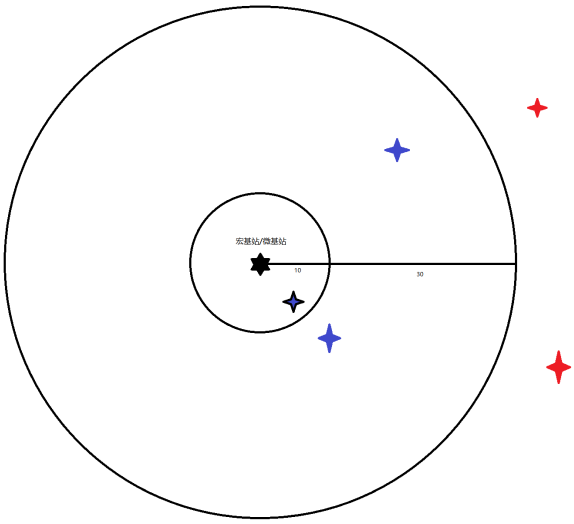

Figure 1. Simulation of base station distribution map

2.2. Problen Two

The second Problem actually adds the restriction of the angle between the base stations to the first Problem. [2] We need to add another three sector models to the first Problem, and need to determine the appropriate angle and make the angle between each two sectors not less than 45 degrees, and consider other conditions such as the base station cost in the previous Problem, and determine the optimal station site and sector angle for the total service volume of the weak coverage points in the range to reach 90% of the overall under the conditions of Problem one.

2.3. Problem Three

Having solved problem one, problem three is actually quite simple and is a less complex data clustering. We need to cluster in regions, requiring 2 weak coverage points with a distance of no more than 20 and considering the condition of transferability.

3. Mathematical Model of Problem One

3.1. Data Acquisition and Pre-processing

The (x,y) coordinates of the base station and the associated service volume can be obtained from the weak coverage raster data (screening) in Annex 2. From the image we can learn that the size of the business volume seen by the stations varies greatly and the level hierarchy is not uniform. And because of the large number of base stations and large data, we need to apply the corresponding mathematical analysis to reduce the data volume to facilitate the calculation.

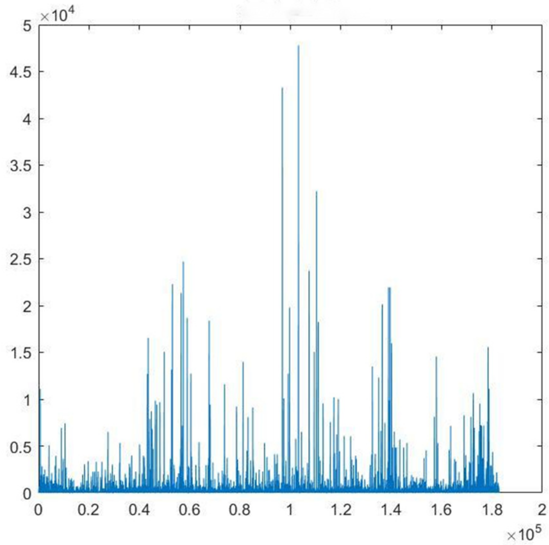

Figure 2. Business volume of each base station

Using simple clustering analysis, a range with a distance of 5 is selected as the clustering criterion, with the purpose of meeting the conditions of coverage construction for macro base stations and micro base stations. Next, the business volume of 99% location is selected as the demarcation to divide the site planning range that will be considered in focus, and then the invalid data including 0 is cleared.

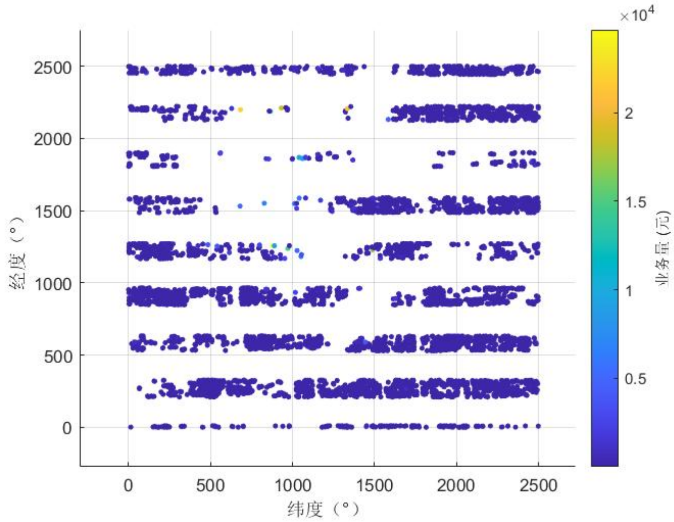

Figure 3. Basic information of weak coverage points

3.2. Model Building

3.2.1. K-Means Clustering Analysis

Means clustering analysis points and groups the objects closest to them. By iterative method, the values of each clustering center are updated one by one until the best clustering result is obtained.

(1) Select c categories as the initial centers for the sample set

(2) In the kth iteration, find the distance of any sample to each center and assign it to the category where the center with the closest distance is located

(3) Calculate the sample centroids of each category as the new center values

(4) If the value of c centers of each cluster remains unchanged after several iterations of the method, the iteration ends, otherwise, the update continues.

ISSN: 2710-0723

3.2.2. One-round Screening

In K-Means cluster analysis, we first define K=500 due to the disorderly nature of the data, and first we sort the points according to the business volume. For the problem of varying business volume and incomplete data consistency, we decided to use K-Means cluster analysis for its initial analysis.

3.2.3. Second-round Screening

Then calculate the distance between all base stations and the location of the station to be considered for establishment, the previous steps for judgment and testing, the previous division process for the priority part of the analysis, to determine the number of base stations to be built in this part. Then continue to use K-Means cluster analysis for the second round of screening, at this time k selected 200, the existing base stations once again to test, remove the existence of error selection, and finally determine the range of base stations to meet the conditions.

3.2.4. Constraints Considered

(1) Minimize the cost.

(2) Threshold of distance between new sites and between new and existing sites ≤ 10.

(3) The coverage area is more than 90 percent.

(4) The coordinates of the site can only be selected among 2500 x 2500 points in the given area.

3.3. Algorithm Implementation

(1) Read the data in Annex I and Annex II.

(2) Sort the data in Annex II according to the business volume.

(3) Use the simple clustering method to determine the priority range using the value of 99% of the business volume as the division boundary.

(4) Verify that the points that exceed the boundary are clear of invalid data including 0.

(5) Calculate the distance of base stations to see if the construction conditions of macro base stations and micro base stations are met.

(6) Based on the previous logic, then build additional base stations for some of the base stations that are not prioritized. Finally, the number of all constructed base stations are combined together for the result test and the cost is calculated.

(7) Then a secondary test is performed, starting from step 5 again.

(8) Finally, the coverage quantities are integrated to get the final results.

3.4. Solution of Problem 1

According to the above steps, our final results have some errors, but the final number of macro base stations is 1125 and 85 micro base stations out of 2500×2500 points in the given area, which is more than 90% coverage.

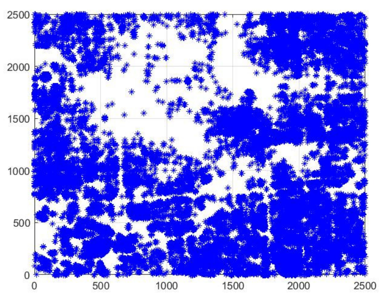

Figure 4. Weak coverage base station locations

ISSN: 2710-0723

Table 1. Problem one solution results

Macro base station required	Micro base station required	Proportion of weak signal points not covered	Percentage of uncovered weak business volume
1125	85	0.0952	0.0986

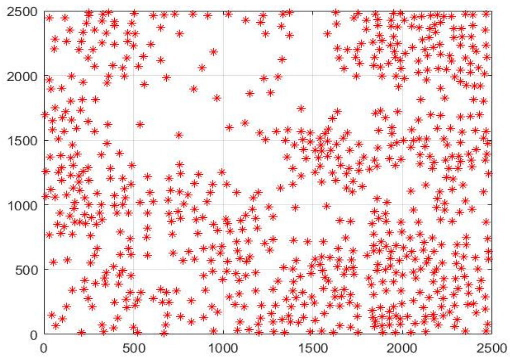

Figure 5. Planning base station sites

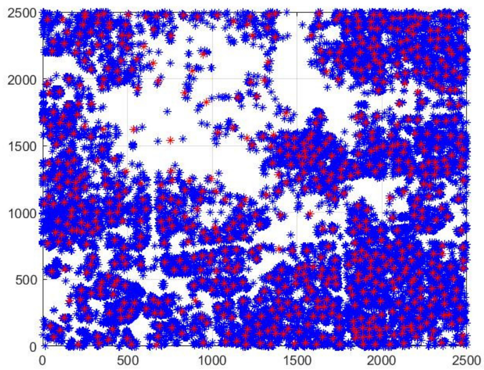

Figure 6. General coverage map

3.5. Analysis of Results

According to the established model and the algorithm to solve problem one, the results of station planning are obtained as shown in Table 1. Among them, 1125 macro base stations are set up, 85 micro base stations are set up, and the total service coverage reaches the standard of 90% with the lowest cost.

4. Mathematical Model of Problem Three

4.1. Model Building

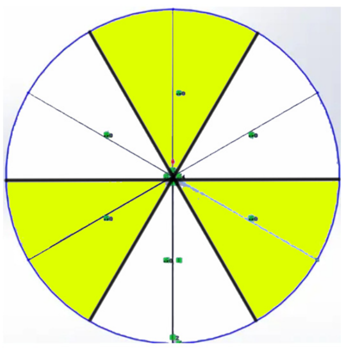

Figure 7. Sector model with three cases

ISSN: 2710-0723

Since Problem 1 mainly addresses the planning of base station coverage, Problem 2 on the basis of Problem 1 only requires the import of three sector models for finding the optimal station site and sector angle that can satisfy the conditions of the optimal station site and sector angle.

4.2. Algorithm Implementation

(1) Read the results of the first Problem and create an empty set to facilitate replication.

(2) Calculate the coordinates of the clustering centers, the distance between two points, and compile the codes separately.

(3) Calculate the angle between base stations separately and perform clustering, compressing the computational effort while considering the case that the clustering points are in different quadrants.

(4) Add three sector models and calculate the sum of the numbers that satisfy the conditions respectively.

(5) Finally, calculate the ratio of the total service volume of weak coverage points that can be covered more.

4.3. Solution of Problem Two

Our three sector models use 30°, 150°, and 270° as the cutoff points to calculate the optimal station and sector angle results, respectively. Due to the complexity of the code and slight errors, we calculate a larger total service satisfying the conditions, which is expressed in scientific notation as 2.182956802195997e+06.

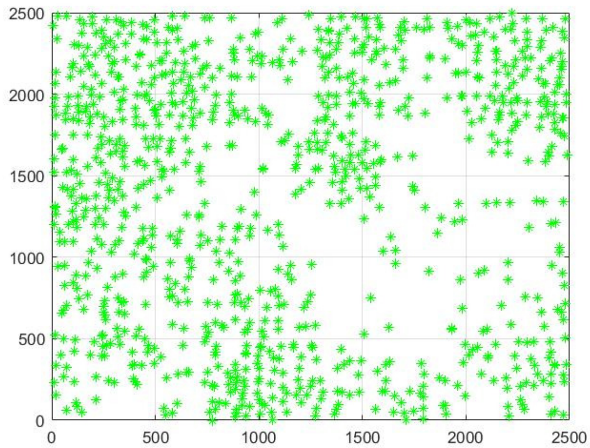

Figure 8. Optimal site distribution map

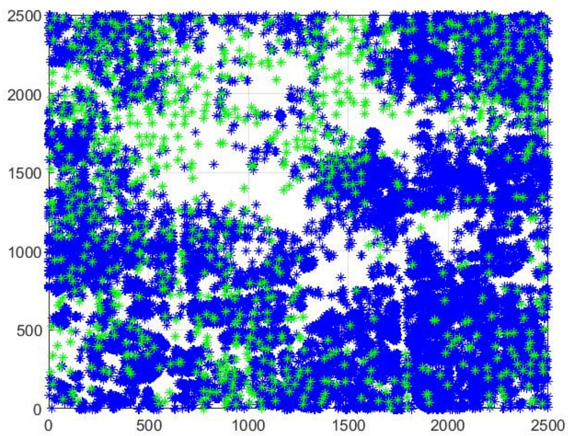

Figure 9. Comparison with initial weak coverage base station sites

ISSN: 2710-0723

4.4. Analysis of Results

The model solution can be obtained that the coverage results are optimal when the selected directions are 30 degrees, 150 degrees and 270 degrees, and the visualization of the results is shown above. When the coverage shape of the base station is changed, it is obviously different from the results of the first Problem, and the total service coverage reaches 90% at this moment, but the number of new base stations can be reduced significantly, which greatly reduces the actual cost.

5. Mathematical Model of Problem Three

5.1. Model Building

The third Problem is similar to the first Problem in that it directly uses clustering to establish clustering relationships based on the condition that the distance between the two weakly covered points is not greater than 20. Here, the K-means clustering method mentioned in the first Problem is mainly used.

5.2. Algorithm Implementation

(1) Derive the weak coverage points from the data obtained in Problem 1, and randomly select them as K clustering centers.

(2) Calculate the distances of the surrounding points to these K cluster centers respectively.

(3) If the distance is not greater than 20, then this coverage point is classified as a class.

5.3. Solution of Problem Three

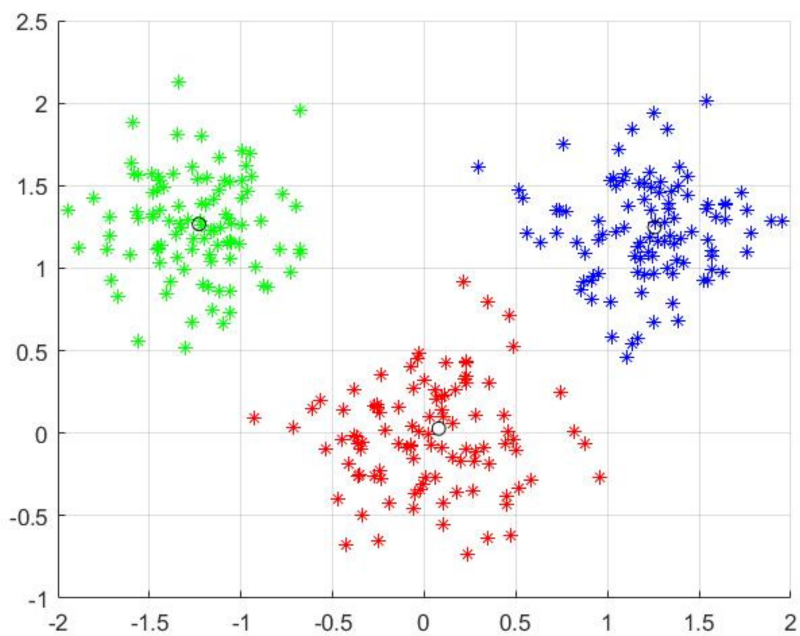

Figure 10. Clustering image of the final result

5.4. Analysis of Results

According to the clustering model results, the final weak coverage area is roughly divided into three categories through continuous clustering iterations, and there is almost no overlap among the three categories, so that the total weak coverage area can be retained to the maximum extent.

ISSN: 2710-0723

6. Evaluation of the Model

6.1. Advantages

(1) Data clustering is used for data cleaning to reduce the amount of computation, while taking into account the relevant conditions given in the Problem to meet the requirements of the Problem as much as possible.

(2) The logic of thinking is clearer, the model is easy to understand and operate, and the algorithm is chosen to solve the problem in a relatively simple and easy to understand way.

6.2. Disadvantages

(1) The calculation method of cost minimization in problem 1 is not optimal.

(2) The sector planning in Problem 2 is cruder, with only 3 types, which has limitations.

(3) The data is too large and the algorithm is not optimized enough, which may require processing more data, thus leading to a slower computing speed.

References

[1] Liu L,Wang P,Pang Zefeng. Agile planning scheme for 5G station sites based on continuous clustering algorithm[J]. Telecommunications Engineering Technology and Standardization, 2022, 35 (03): 68-71..2022.

[2] Liu Yaxi. Research on coverage calculation and optimization methods for mobile communication networks. University of Science and Technology Beijing.2021.

[3] Sun Peng,Yang Dacheng. Research on optimization algorithm for wireless network planning[J]. Radio Engineering,2005(03).