

Article

Application of Clustering Algorithms in the Location of Electric Taxi Charging Stations

Qing Li ¹, Xue Li ¹, Zuyu Liu ¹ and Yaping Qi ^{2,*}

¹ College of Mathematics and System, Shandong University of Science and Technology, Qingdao 266590, China; i_liqing@163.com (Q.L.); lixuea2021@163.com (X.L.); liuzuyu2021@163.com (Z.L.)

² College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao 266590, China

* Correspondence: qiyaping79@126.com

Abstract: The reasonable layout of charging stations is an important measure to improve the penetration rate of the electric taxi market. Based on the multi-type clustering algorithm, a widely applicable electric taxi charging stations locating method is proposed. By analyzing the massive gasoline taxi GPS trajectory data, the parking information and charging requirements of electric taxis are extracted, and the research area is divided into reasonable grids. Then, the divided grids are respectively subjected to multiple same-type clustering and multiple multi-type clustering algorithms, so as to help find out the location of the charging station, and a comparative analysis is performed. The empirical analysis shows that the positioning results of the multiple multi-type clustering algorithms are more reasonable than the multiple same-type clustering algorithms, which can effectively prolong the driving distance of electric taxis and save the travel time of drivers.

Keywords: charging station location; multiple and multi-type clustering; electric vehicle; GPS trajectory data



Citation: Li, Q.; Li, X.; Liu, Z.; Qi, Y. Application of Clustering Algorithms in the Location of Electric Taxi Charging Stations. *Sustainability* **2022**, *14*, 7566. <https://doi.org/10.3390/su14137566>

Academic Editors: Ruifeng Shi, Li Ji, Han Liu and Zhe Zhang

Received: 30 April 2022

Accepted: 1 June 2022

Published: 21 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rapid development of the global economy and the continuous reduction of fossil energy [1], the greenhouse effect is becoming more and more serious. Low-pollution, low-emission electric vehicles (EVs) are gradually attracting great attention [2]. Batteries are the main power source of EVs, but due to the limitations of current battery technology, EVs are not practical for long-distance travel and charging. This can lead to mental distress or apprehension caused by the driver's fear of suddenly running out of power when driving an electric vehicle (range anxiety) [3], which has become an important factor hindering the development of EVs. It can be concluded from this that the key to solving the driving range problem of electric vehicles is to optimize the layout of charging stations according to the charging demand.

The reasonable deployment of charging infrastructure plays a positive role in extending the driving range of EVs and promoting the development of EVs. In order to solve the problems of insufficient range, inconvenient charging, and unreasonable charging infrastructure layout of EVs, a lot of studies have investigated the location problem of public charging stations for EVs from different perspectives.

In terms of the influencing factors, some researchers have found that charging demand [4], vehicle miles traveled [5,6], geographic distribution of cities [6], path deviation [7], traffic flow patterns [8], and other factors are important in influencing the location of charging stations, of which charging demand is the most fundamental factor. The current market penetration of EVs is low, and the accurate estimation of charging demand can help achieve an optimized layout of charging stations.

To ensure more accurate estimation of charging demand, some researchers have used GPS trajectory data of gasoline vehicles to simulate the trajectory data of EVs for charging infrastructure location selection. For example, Pan et al. [9] used survey data of household trips

to simulate the charging selection behavior of drivers with the decision process of EV charging selection so that the existing travel activities of drivers are maximally unaffected. Chen et al. [10] used parking information from more than 30,000 individual trip records collected from a household trip survey in Seattle, Washington, USA, to determine the optimal number of charging stations to be allocated. Liu et al. [11] proposed an intelligent optimization method as well as data-driven and particle swarm optimization based on GPS trajectory data of hybrid vehicles in Chengdu, China, to achieve intelligent siting of EV charging stations. Yang et al. [12] used the GPS trajectory information of a fleet of cabs in Changsha, China, to estimate the likelihood of EV charging using a queuing model, and investigated the relationship between installing more charging piles and the trade-off between providing more waiting space, and the effect of charging power on waiting time. Due to the short time of the emergence of electric taxis (ETs) in China, relevant GPS trajectory data are scarce to give a reasonable layout of charging stations. Shi et al. [13] present an improved destination selection model, proposed to simulate the ET operation system and to help find the optimal ET charging station size with statistical analysis based on the charging need prediction. Therefore, this paper simulates the travel trajectory of electric taxis with GPS trajectory data of gasoline vehicles and predicts the charging demand of electric taxis by combining this with grid-based maps.

In terms of location methods, most studies focus on constructing charging station location models and siting charging stations based on different objective functions and constraints [14,15], while relatively few studies have applied clustering algorithms to the location of electric vehicle charging stations. Cluster analysis is a kind of unsupervised learning, and there are many kinds, such as the division-based K-means clustering algorithm [16], hierarchical clustering algorithm (agglomerative and splitting) [17], and DBSCAN algorithm based on density clustering, etc. At present, clustering algorithms are widely used in short-time traffic flow prediction [18], logistics center location selection [19], traffic flow speed prediction [20], and travel hotspot area research [21], etc., but they are less widely applied in charging station location selection. For example, Zhang et al. [22] developed a siting model for electric cabs based on their dynamic distribution and charging demand using the K-means clustering method and the center of gravity method, and applied it to the problem of siting electric cabs in Chengdu, China. Straka et al. [23] analyzed charging transactions in the Netherlands using clustering algorithms (K-means, dbSCAN, and cohesive hierarchical clustering) to identify usage related segments of charging stations, which helps to improve the planning of charging infrastructure and the development of smart charging technologies. Liu et al. [24] used existing service areas on highways as potential locations for charging infrastructure, clustered the close service areas, and calculated the optimal location of charging stations for each cluster. Gilanifar et al. [25] proposed a Gaussian process based on the Clustered Multi-Node Learning (CMNL-GP) method to fuse and learn data from multiple charging stations simultaneously. Zhang et al. [26] proposed a density peak clustering-based optimization method for siting and sizing EV charging stations in an urban area. Sánchez et al. [27] proposed a clustering strategy based on the K-means algorithm to define potential charging station locations. The above studies are based on different clustering algorithms for the siting and sizing determination of charging stations, and there are not yet multiple clustering algorithms combined and applied in the siting study.

This paper uses multiple and multi-type clustering algorithms to optimize the location of charging stations and obtain the optimal charging station location and clustering algorithm combination. At present, electric taxis have not been fully popularized in Qingdao. Therefore, we can only estimate the charging demand of electric taxis through the GPS trajectory data of gasoline taxis in reality. The research areas are five main districts of Qingdao (Shinan District, Shibei District, Licang District, Chengyang District, and Laoshan District). Firstly, the map of the study area is gridded, and the number of vehicles in each grid that stay longer than the time threshold is recorded as the number of dwell events, and the number of dwell events in each grid is used as the charging demand of the grid. Finally, the overall weighted Euclidean distance sum of the two location selection methods is compared, and the optimal location and the best location selection method for charging stations are obtained. The proposed method is of theoretical and practical significance as it

provides quantitative charging station location solutions to help deal with actual charging station location problems in urban management.

This paper is organized as follows: Section 2 presents the problem statement and data processing. Section 3 presents the location selection methodology. Section 4 presents the results of the charging station location. Finally, we summarize this paper and present the limitations and future research directions in Section 5.

2. Problem Statement and Data Processing

2.1. Problem Statement

Given a set of electric vehicles that charge at least once a day at a charging station and a set of grids with a long dwell of electric vehicles, the problem is to present multiple and multi-type clustering algorithms to optimize the location of charging stations, so as to obtain the optimal charging station layout and the best clustering siting algorithm combination.

2.2. Data Processing

In this paper, by analyzing the taxi GPS trajectory data from 0:00 to 24:00 on 11 October 2017 (a weekday) and combining the economic development and geographical location characteristics of each urban area in Qingdao, it is extracted from the taxi GPS trajectory data within five major municipal districts. By removing the abnormal data and discontinuous trajectory data, we finally obtained 828,341 GPS trajectory data for a total of 6042 taxis. The GPS data of each taxi was captured approximately once per 5 s. Data items include vehicle ID, time, longitude, latitude, speed (km/h), and passenger status (0 for empty, 1 for passenger) (See Table 1).

Table 1. The examples of GPS trajectory data.

Vehicle ID	Time	Longitude	Latitude	Speed	Passenger Status
30	18:52:44	120.300572	36.059037	7.4	0
600	18:48:08	120.299342	36.05908	22.4	0
850	18:52:47	120.338533	36.05844	40.8	1

The research area of this paper consists of the five main municipal districts of Qingdao (Figure 1). In order to exhaustively count the dwell demand in each area and prevent incorrect statistics of dwell demand due to inconsistent area size, this paper divides the research area into 4759 grids according to the size of both image width and height of 0.005° (about a rectangular grid of $450\text{ m} \times 550\text{ m}$), and automatically numbers them as i ($i = 1, 2, \dots, 4759$), so the size of the grid number is only for indication, and cannot be used to indicate the distance of geographical location between the grids. Therefore, the geometric center of the grid is chosen to represent the geographic location of the grid, and the number of dwell events of the grid represents the charging demand of electric taxis in the grid.



Figure 1. Five main municipal districts of Qingdao. Note: The map data used in this article comes from Gaode Map, the same below.

2.3. Charging Demand

In order to maximize the satisfaction of the charging demand of electric taxis and enhance the rationality of site selection, this paper assumes that the travel patterns of drivers will not change during the electrification of gasoline taxis. So the GPS travel trajectory data of gasoline taxis in the five main districts of Qingdao is used to simulate the travel behavior of electric taxis, and extract the vehicle parking patterns to mine their charging needs.

Taxi drivers typically have long dwelling times for meals, fuel, shift changes, or breaks, so it is reasonable to assume that electric taxis will have charging needs during this time. In order to fully understand the charging demand of taxi drivers, this paper uses a time threshold of 20 min [12] to distinguish the dwelling of vehicles. If the GPS trajectory data shows that the vehicle dwell in the same grid for more than 20 min, it is considered that the vehicle needs to be charged in this grid, and a dwell event occurs in this grid.

The research area in this paper was divided into a total of 4759 grids, and the total number of dwell events occurring in each grid was recorded. Of the 4759 grids with dwell events, 900 occurred in a total of 3202 taxis, with 3401 dwell events. Figure 2 shows the frequency statistics of the number of dwell events per grid. It can be seen from Figure 2 that the number of grids that did not have a dwell event accounted for 81% of the total number of grids, the number of grids that had a stay event accounted for 6%, and the number of grids with 20 or more dwell events is the least.

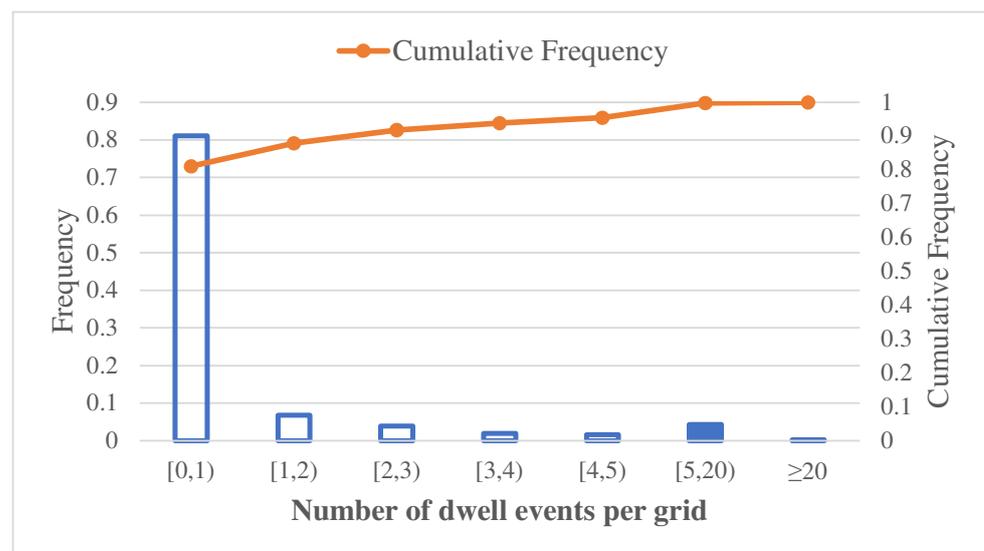


Figure 2. Frequency distribution of the number of dwell events per grid.

Figure 3 shows the number of dwell sites at which one driver would dwell in a day. Most taxis only have one dwell event in a day, and a taxi can stay in up to three locations for more than 20 min in a day. Figure 4 shows the spatial distribution of dwell events (a represents the number of dwell events per grid). As seen in Figure 4, the grids with dwell events are relatively dense, and the grids with many dwell events are mostly located in the center of each district. The eastern part of Laoshan District is Laoshan Scenic Area, and the electric vehicle dwells are relatively few and scattered, so there are fewer grids with dwell events; the development of the eastern and western parts of Chengyang District is more different, and the western area is relatively backward and has fewer electric vehicles, so the grid distribution in Figure 4 is consistent with the actual situation of EV dwells in the city.

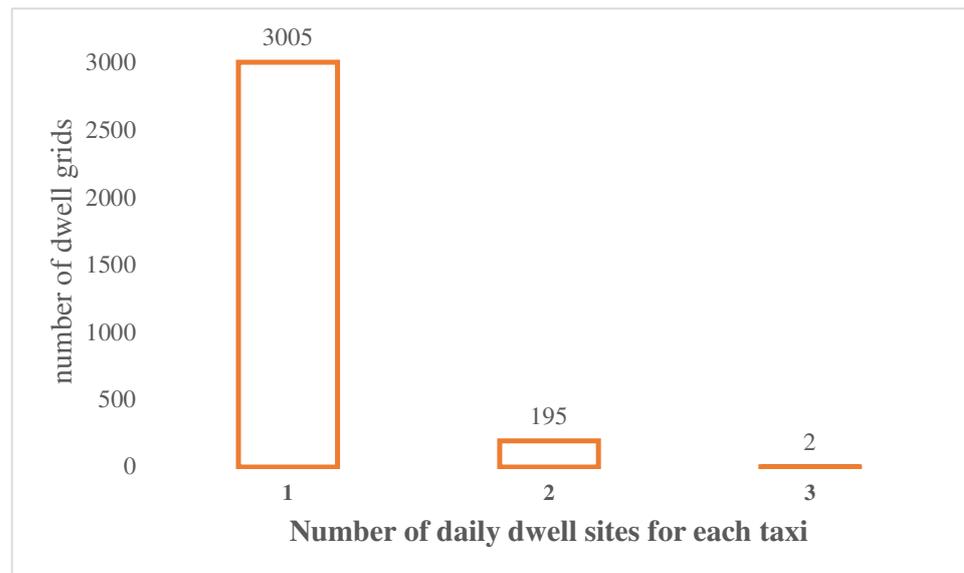


Figure 3. Distribution of the number of dwell sites for each taxi.

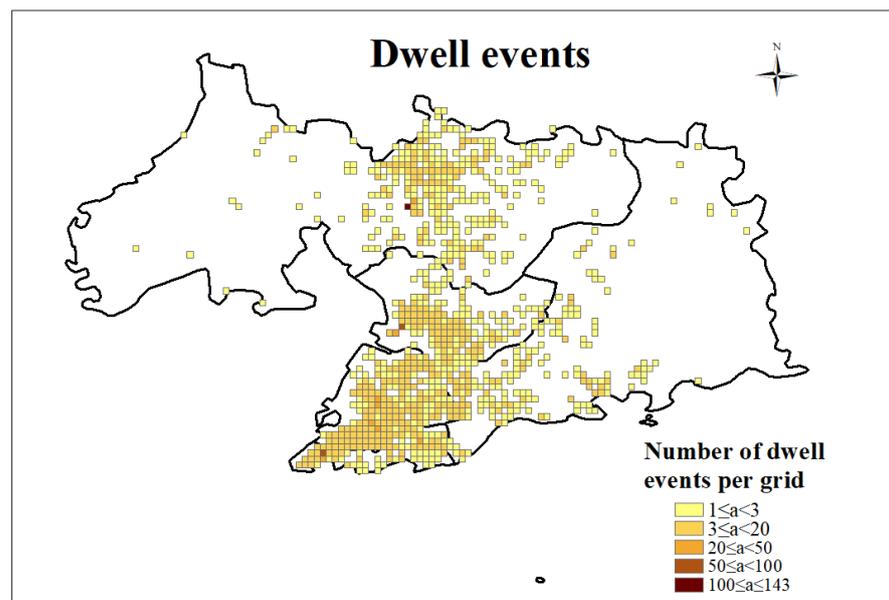


Figure 4. Distribution of dwell events.

In this paper, the number of dwell events in each grid is used to represent the charging demand for electric taxis in the grid, and the higher the number of dwell events occurring, the higher the charging demand for EVs in the grid. It is relatively uneconomical to install charging stations in places that are not attractive to taxi drivers. Therefore, grids with no less than 4 ($a \geq 4$) number of dwell events are selected as the study object, and it can be found that 295 grids out of 4759 grids satisfy the condition and contain 2312 taxis. Therefore, this paper optimizes the location of charging stations based on the 295 grids where taxis dwell.

3. Methodology

To achieve the optimal layout of charging stations and find the best combination of clustering sizing algorithms, this paper proposes the multiple and multi-type clustering algorithms, which mainly involve the K-means clustering algorithm, K-means weighted clustering algorithm, and hierarchical clustering algorithm.

3.1. Calculation of Euclidean Distance

The difference between the K-means clustering algorithm and the K-means weighted clustering algorithm is whether Euclidean distance or the weighted Euclidean distance is used in the clustering process.

The Euclidean distance calculation formula of the K-means clustering algorithm is

$$D(x_i, x_j) = \sqrt{\sum_{a=1}^m (x_{ia} - x_{ja})^2} \quad (1)$$

where $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$, $x_j = (x_{j1}, x_{j2}, \dots, x_{jm})$ represent two data objects containing m -dimensional attributes [15].

In this paper, the position coordinates of the i -th grid can be expressed as

$$L_i = (xpos_i, ypos_i), i \in N \quad (2)$$

where L_i represents the position of the i -th grid; $xpos_i$ and $ypos_i$ denote the x-coordinate and y-coordinate of the i -th grid position expressed in terms of GPS longitude and latitude, respectively. N is the set of grid numbers in this category.

Suppose the coordinates of the k -th cluster center can be written by Equation (3)

$$Z_k = (zx_k, zy_k), k = 1, 2, \dots, K \quad (3)$$

where Z_k represents the position of the k -th cluster center; zx_k and zy_k represent the x-coordinate and y-coordinate of the k -th cluster center, respectively.

Therefore, the calculation formula of the Euclidean distance between the i -th grid and the k -th cluster center can be written by Equation (4)

$$D(L_i, Z_k) = \sqrt{(xpos_i - zx_k)^2 + (ypos_i - zy_k)^2}, i \in N, k = 1, 2, \dots, K \quad (4)$$

In the K-means clustering algorithm, the commonly used methods for determining the number of clusters K are the silhouette coefficient method and the elbow rule. The silhouette coefficient method determines the optimal K value by finding the local optimal result; the elbow rule determines the optimal K value by judging the change of the sum of squared errors (SSE) within the class.

This paper uses the elbow rule to determine the number of clusters K . In the elbow rule, the sum of squared errors (SSE) of the distance between the cluster center of each class and the sample points in the class is called the degree of distortion. For a class, the lower the degree of distortion, the closer the sample points within the class are. The more the number of clusters, the fewer sample points each class contains, and the closer the sample points are to the center point of the cluster, so the degree of distortion will decrease with the increase of the number of clusters. If the number of clusters exceeds the actual number of categories, the degree of distortion changes little, even if the number of clusters K increases, the degree of distortion does not change significantly, so an area similar to "elbow" will be formed on the line graph composed of the degree of distortion and the corresponding K value of the elbow is the selected number of clusters. The formula for calculating the degree of distortion (SSE) can be written by Equation (5)

$$SSE = \sum_{k=1}^K \sum_{i \in N} D(L_i, Z_k)^2 \quad (5)$$

In addition, the K-means weighted clustering algorithm selects a weight in the K-means clustering process and improves the Euclidean distance into the weighted Euclidean

distance, so the weighted Euclidean distance between the position of the i -th grid and the center of the k -th cluster is calculated (Equation (6)).

$$D(L_i, Z_k) = \sqrt{w_i((xpos_i - zx_k)^2 + (ypos_i - zy_k)^2)}, i \in N, k = 1, 2, \dots, K \quad (6)$$

where w_i is the weight of the i -th grid.

3.2. Multiple Same-Type Clustering and Multiple Multi-Type Clustering Algorithms

This paper presents multiple and multi-type clustering algorithms for the siting layout of charging stations. The so-called multiple, that is, repeatedly applying the same clustering algorithm and improving this algorithm in the application; the so-called multi-type, that is, comprehensive application of multiple clustering algorithms.

Method 1: The multiple same-type clustering algorithms first use K-means clustering to obtain the classification results based on the geographical location between grids; secondly, using the charging demand of each grid as the weight, K-means weighted clustering is performed on the sample points of each category. The new cluster center of each class is obtained, which is the location of the charging station, and the intra-class weighted Euclidean distance sum from the sample points of each class to the cluster center is calculated, and finally, the overall weighted Euclidean distance sum is obtained.

Method 2: The multiple multi-type clustering algorithms use K-means clustering to obtain classification results based on the geographical location between grids; secondly, considering the charging demand of each grid, the two-step clustering method is used to select the location of charging stations. The so-called two-step clustering method is to perform agglomerative hierarchical clustering for each class of sample points, choose a fixed relative distance to reclassify each class of sample points, and then perform K-means weighted clustering for each class of sample points after classification to obtain the location of charging stations, and calculate the intra-class weighted Euclidean distance sum for each class. Therefore, finally, we obtain the overall weighted Euclidean distance sum.

4. Results

In order to obtain a reasonable layout and siting method for charging stations, this paper firstly clusters the screened 295 grids using a K-means clustering algorithm based on the geographic location attributes between the grids. The input data samples of the K-means clustering algorithm are shown in Table 2, which contains the grid number, the longitude and latitude corresponding to the grid location, and the number of dwell events for the grid.

Table 2. The examples of the input data for K-means clustering algorithm.

Grid Number	Longitude	Latitude	Number of Dwell Events
22	120.2984848	36.056797	4
282	120.3634796	36.0967979	16
434	120.3484802	36.1117973	14
3752	120.3984833	36.2917976	7

The choice of K value is crucial for K-means clustering algorithm. Table 3 shows the number of clusters K and the specific values of the corresponding degree of distortion (SSE) obtained by the elbow rule, and Figure 5 is the elbow diagram obtained by the elbow rule. According to the elbow diagram judgment K value is obtained by human subjective observation, so this paper sets a limit value for the variation difference of the degree of distortion (SSE). If the variation difference of the degree of distortion is less than this limit value, the former K value is selected. The limit value of the variation difference of the degree of distortion is set to 0.02, so K can be set to 4. Meanwhile, through the observation of the elbow diagram in Figure 5, it can be seen that when $K > 4$, the degree of distortion

(SSE) does not change significantly, so the final number of clusters K is determined to be 4. The simultaneous use of the two methods ensures the accuracy of the K value.

Table 3. The number of clustering and the corresponding distortion degree.

Number of Clusters K	Degree of Distortion (SSE)
1	0.554661
2	0.231372
3	0.121694
4	0.065528
5	0.041640
6	0.028624
7	0.021175
8	0.016797
9	0.012739

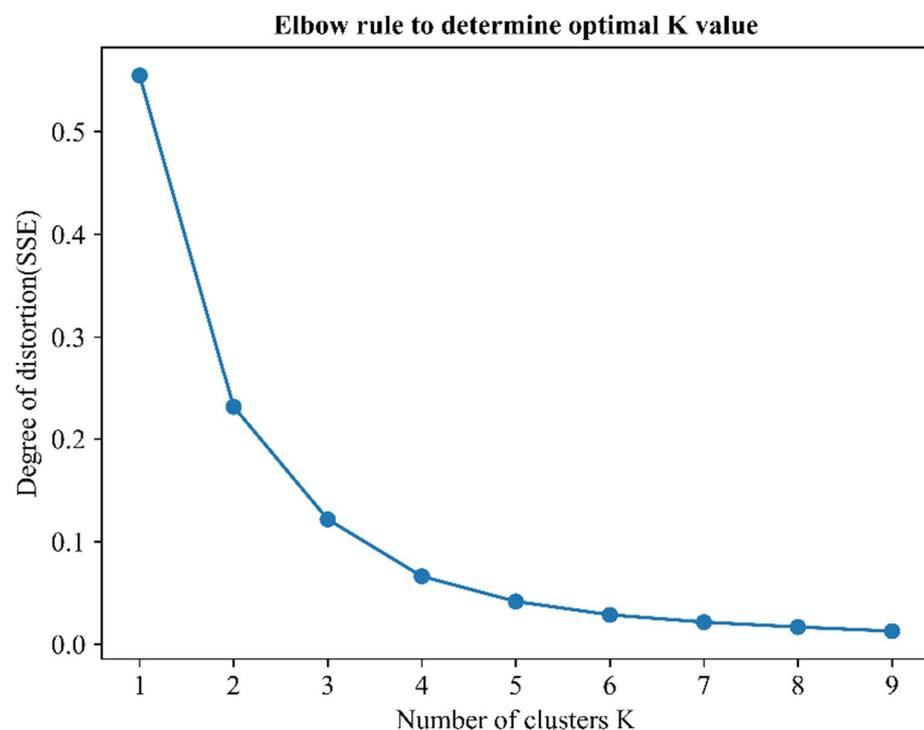


Figure 5. Elbow figure.

Figure 6 shows the clustering diagram of K-means clustering with selected $K = 4$. The dots of different colors represent different clustering categories. The geographical locations between the grids within each category are the closest, and the number of grids in the 4 categories are 118 (category 1), 21 (category 2), 102 (category 3), and 54 (category 4), respectively. From Figure 6, we can see that the sample points in category 1, category 3 and category 4 are more concentrated, while the sample points in category 2 are more dispersed. The grids in category 2, belonging to the Laoshan District, are unevenly distributed because there is a large mountainous area in Laoshan District.

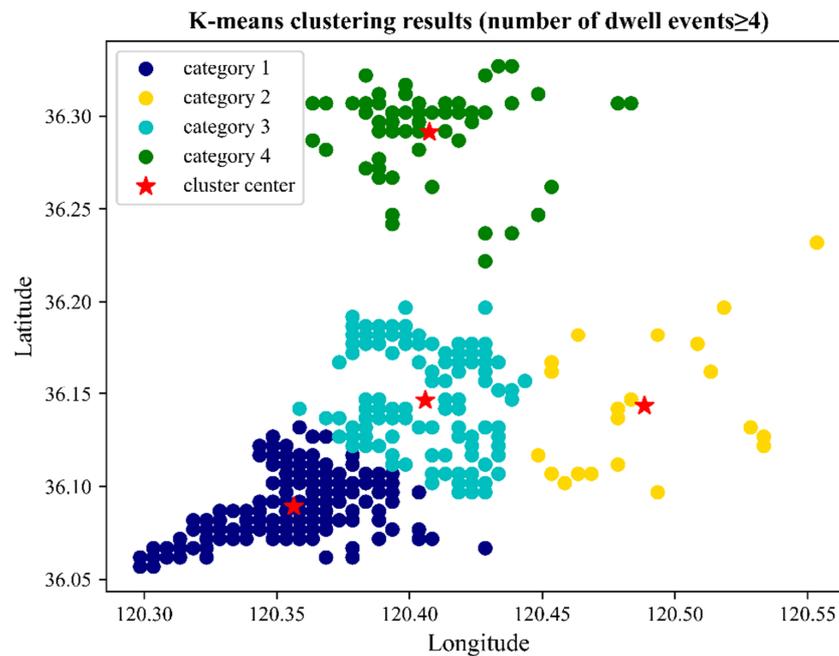


Figure 6. K-means clustering results.

K-means clustering only considers the geographic location between 295 grids without considering the charging demand of electric taxis in each grid, so the 4 clustering centers obtained are not the best locations for charging stations. Therefore, the following two optimization modes of location selection are chosen to select the optimal location of the charging station on the basis of the K-means clustering results.

4.1. Location Results of Multiple Same-Type Clustering Algorithms

The more the number of dwell events in the grid, the greater the charging demand of electric taxis in the grid. According to the 4 classes of grid data obtained by K-means clustering, the charging demand of electric taxis in each grid is considered, and the charging demand of electric taxis in each grid is used as the weight of the grid. Then K-means weighted clustering ($K = 1$) is performed for each class of the sample points to obtain 4 new cluster centers and the corresponding intra-class weighted Euclidean distance sums (Table 4), and finally obtain the overall weighted Euclidean distance sum of the four types of grid data is 24.1.

Table 4. The cluster centers for K-means weighted clustering algorithm.

Number	Longitude	Latitude	Intra-Class Weighted Euclidean Distance Sum
Cluster Center 1	120.35636442	36.08921391	9.314467822
Cluster Center 2	120.48848397	36.14346568	1.946304218
Cluster Center 3	120.40593398	36.14640666	8.401443127
Cluster Center 4	120.40746448	36.29133549	4.439257816

Figure 7 shows the best locations of charging stations obtained by the multiple same-type clustering algorithms. As can be seen from Figure 7, charging stations are located in the economic and residential centers of Shibei District, Licang District, Laoshan District, and Chengyang District, respectively. However, the small number of charging stations may cause long queues of electric taxis and reduce drivers' satisfaction with charging, while the long queues may cause traffic congestion problems around the charging stations. Therefore, it is necessary to further deal with.

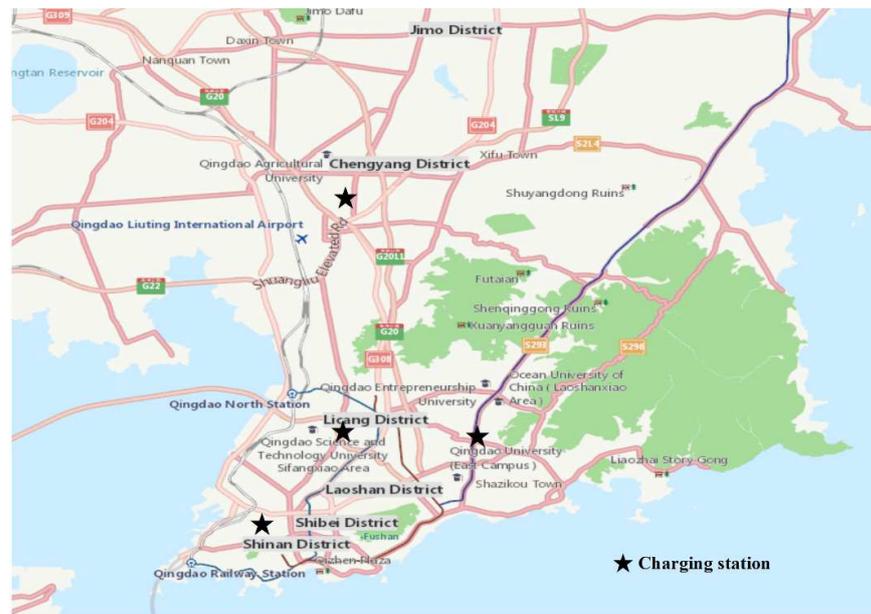


Figure 7. The layout of charging station (Multiple Same-type Clustering Algorithms).

4.2. Location Results of Multiple Multi-Type Clustering Algorithms

Based on the results of K-means clustering, a two-step clustering method (agglomerative hierarchical clustering and K-means weighted clustering) is used to optimize the location of charging stations. Firstly, agglomerative hierarchical clustering is performed on the sample grid data of each class obtained by K-means clustering, and a relative distance of 0.08 is selected to classify the grid. Figure 8 shows the tree diagram obtained by agglomerative hierarchical clustering for each of the 4 classes of sample grid data. The black dotted line in the tree diagram represents the relative height of 0.08, which is used to divide the results of agglomerative hierarchical clustering.

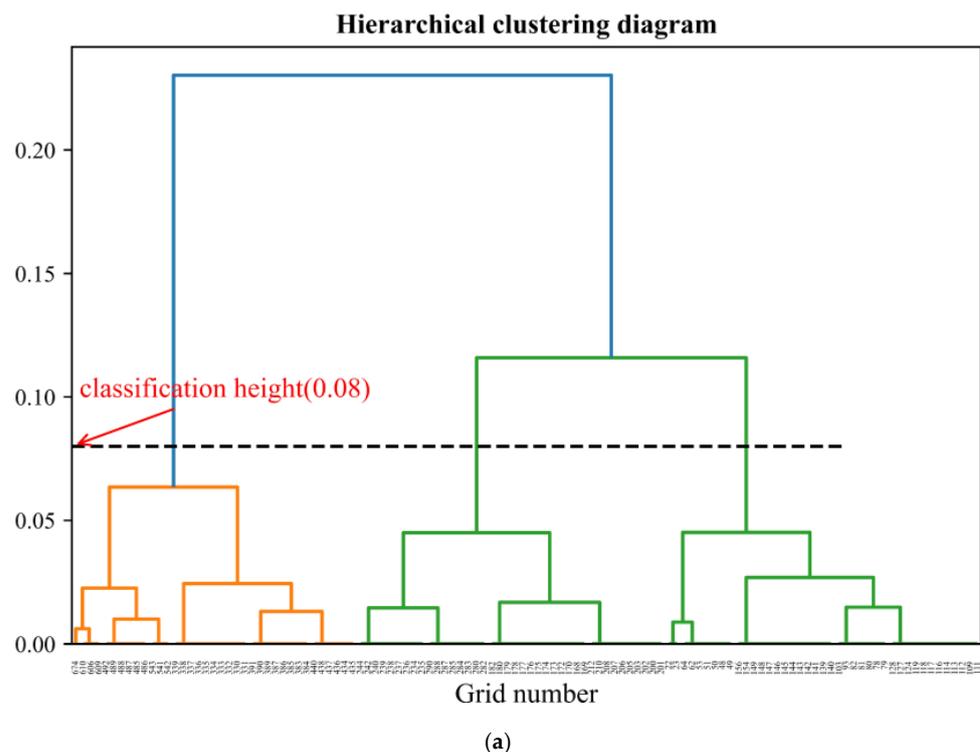
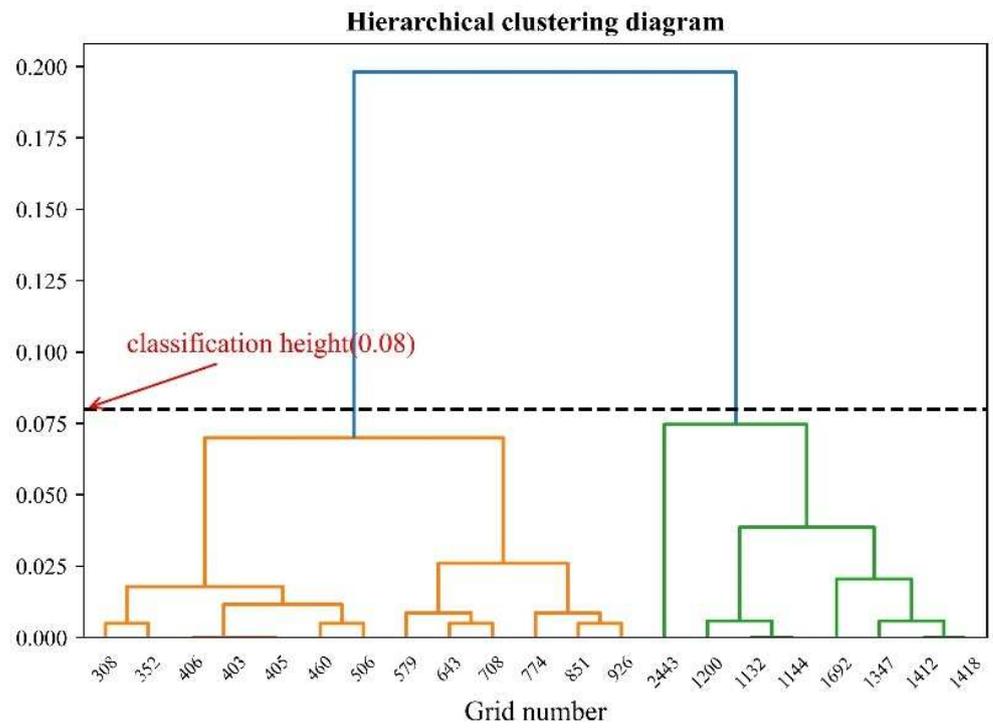
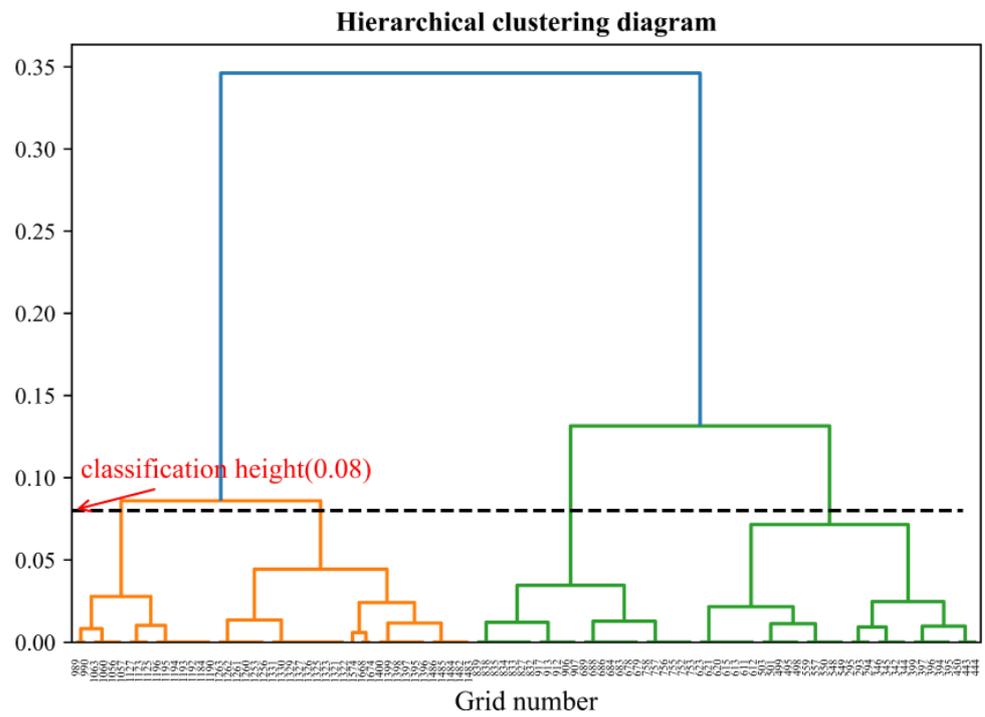


Figure 8. Cont.



(b)



(c)

Figure 8. Cont.

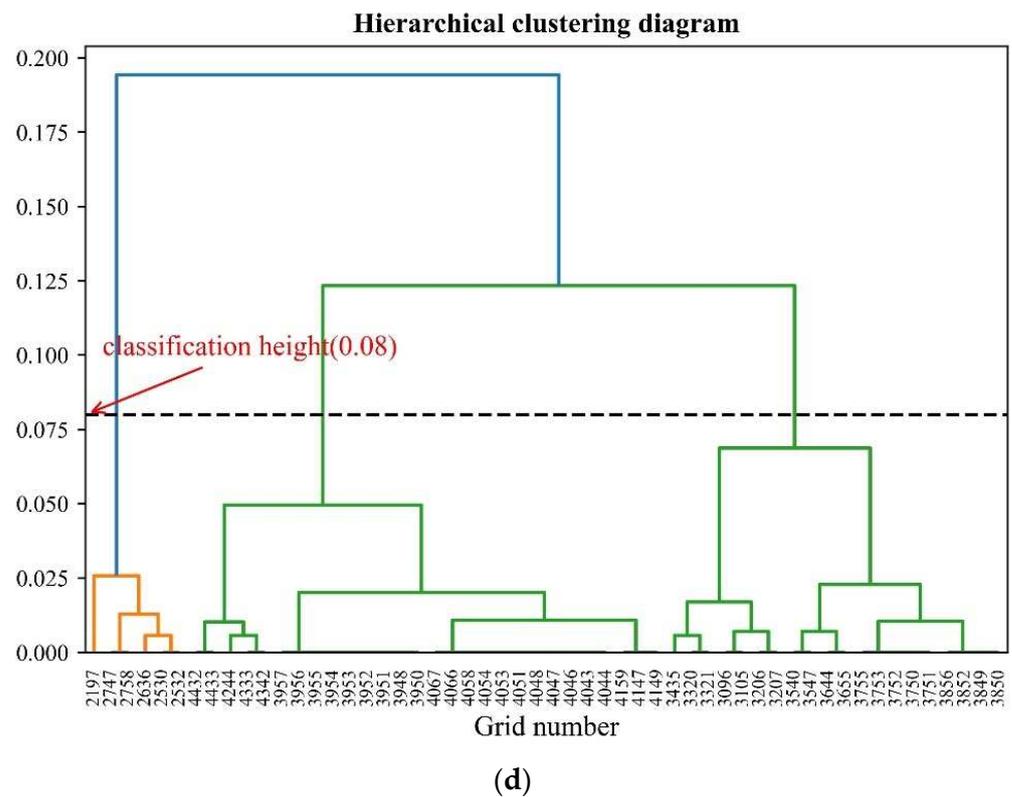


Figure 8. Tree diagram of hierarchical clustering algorithm. (a) Hierarchical clustering diagram for category 1. (b) Hierarchical clustering diagram for category 2. (c) Hierarchical clustering diagram for category 3. (d) Hierarchical clustering diagram for category 4.

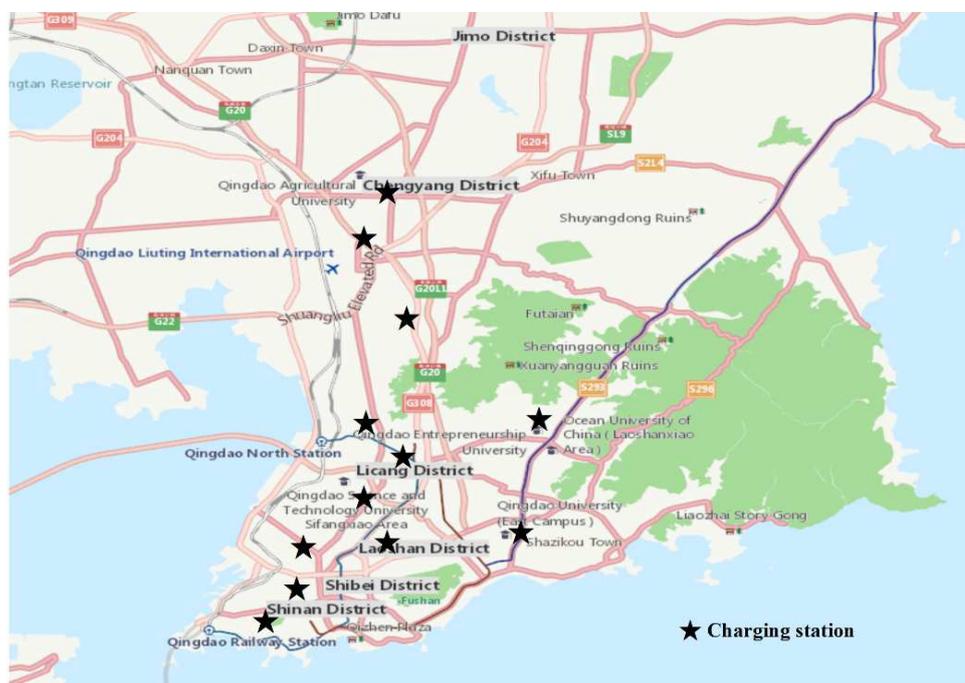
Figure 8a shows the agglomerative hierarchical clustering result of category 1. The grids in the sample are basically located in the southern parts of Shinan District and Shibei District. The areas of the two urban areas are relatively small, but the daily vehicle flow is large and the number of grids is large, so it is divided into 3 categories. Figure 8b shows the agglomerative hierarchical clustering result of category 2. Most of the grids in the sample are located in Laoshan District, and the distance between grids is long, so it is divided into 2 categories. Figure 8c shows the agglomerative hierarchical clustering results of category 3. The grids in the sample are mostly located in the northern part of Licang District and Shibei District, the location between grids is close and the number of grids is large, which indicates that the charging demand of electric vehicles in this area is large. So it is divided into 4 categories. Figure 8d shows the agglomerative hierarchical clustering result of category 4. The grids in the sample are basically located in Chengyang District. Most grids are densely located, and some grids are scattered around, so it is divided into 3 categories. After performing agglomerative hierarchical clustering on grid samples, all sample grid data are divided into 12 categories.

The 12 categories of data obtained by the agglomerative hierarchical clustering method are respectively subjected to K-means weighted clustering ($K = 1$) to obtain 12 cluster centers, which are the optimal locations of charging stations. The 12 cluster centers and the corresponding intra-class weighted Euclidean distance sums are shown in Table 5, and the final overall weighted Euclidean distance sum is 16.1.

Table 5. The cluster centers for Two-step clustering method.

Category	Cluster Center	Longitude	Latitude	Intra-Class Weighted Euclidean Distance Sum
Category 1	Cluster Center 1	120.34409294	36.07009134	3.469152132
	Cluster Center 2	120.36123313	36.08829833	1.945626697
	Cluster Center 3	120.36469908	36.11139359	1.501172971
Category 2	Cluster Center 1	120.48463792	36.11949128	0.839284857
	Cluster Center 2	120.49473381	36.18242406	0.606927985
Category 3	Cluster Center 1	120.41098308	36.11398672	1.574992016
	Cluster Center 2	120.39828278	36.13879853	1.09009133
	Cluster Center 3	120.41973210	36.16148613	1.234712563
	Cluster Center 4	120.39934566	36.18041940	0.714346782
Category 4	Cluster Center 1	120.42181778	36.23846498	0.30493468
	Cluster Center 2	120.39823263	36.28304862	1.135222951
	Cluster Center 3	120.41098295	36.30858407	1.712925032

Figure 9 shows the optimal layout of charging stations obtained by the two-step clustering method. It can be seen from Figure 9 that the location of the charging station matches the grid with many dwell events, which meets the charging demand of electric vehicles. Compared to Laoshan District and Chengyang District, the number of charging stations in Shinan District, Shibe District, and Licang District is higher. This is because the three areas are densely populated with residential areas, commercial areas, scenic spots, and high population density, resulting in high traffic flow, many dwell events, and a high charging demand for electric vehicles.

**Figure 9.** The layout of charging station (Multiple Multi-type Clustering Algorithms).

4.3. Results Analysis

By comparing the overall weighted Euclidean distance sum obtained by the multiple same-type clustering algorithms and the multiple multi-type clustering algorithms, it can be seen that the multiple multi-type clustering algorithms effectively reduces the overall weighted Euclidean distance sum. That is, they reduce the traveling distance from the electric vehicle to the charging station, save the travel time of the electric vehicle driver,

and increase the operating time. From Figures 7 and 9, it can be found that the location and layout of the charging station in Figure 9 is more reasonable, which can meet the charging demand of electric vehicles as much as possible and achieve the goal of optimizing the location of the charging station. Meanwhile, it can be seen that the multiple multi-type algorithms proposed in this paper are better than the multiple same-type clustering algorithms in the application of charging station locations, which provides a new method for future charging station locations.

5. Conclusions

This paper takes the gridded map of five major municipal districts of Qingdao (Shinan District, Shibei District, Licang District, Chengyang District, and Laoshan District) as the research area, extracts the number of vehicles in each grid with a dwell time of more than 20 min, based on the GPS trajectory data of gasoline taxis in the five main municipal districts, and takes the grid with the number of dwell events no less than 4. Assuming that the geometric center of each grid and the number of dwell events, respectively, represent the location of the grid and the charging demand of electric taxis. Then, the location of charging stations is selected using the clustering method. Based on the geographic location among grids, multiple same-type clustering algorithms and multiple multi-type clustering algorithms are performed for all grids separately. Finally, the overall intra-class weighted Euclidean distance sum obtained by the multiple same-type clustering method is 24.1, and the overall intra-class weighted Euclidean distance sum obtained by the multiple multi-type clustering method is 16.1, which indicates that the overall weighted Euclidean distance sum obtained by the multiple multi-type clustering algorithms is significantly smaller than that of the multiple same-type clustering algorithms, reducing the traveling time of electric taxis. The location selection result of the multiple multi-type clustering algorithms is more reasonable than that of the multiple same-type clustering algorithms. This paper provides feasible suggestions and methods for the location and optimal layout of charging stations in five major municipal districts of Qingdao.

Currently, the market penetration rate of electric vehicles is increasing, and the reasonable layout of charging stations plays a positive role in the promotion of electric vehicles. The multiple multi-type clustering location selection method proposed in this paper provides a new solution for the optimal layout of urban charging stations. However, only the travel time of electric vehicle drivers is considered, and the trajectory data of gasoline taxis are used to simulate the trajectory of electric taxis, which has some errors in terms of station location. In future research, the cost problem can be considered, more clustering algorithms can be integrated, and GPS trajectory data of electric vehicles can be used to further improve the scientific and reasonable location of charging stations.

Author Contributions: Conceptualization, Q.L. and X.L.; methodology, X.L.; software, X.L.; validation, X.L. and Z.L.; formal analysis, Q.L.; investigation, X.L.; resources, Q.L.; data curation, X.L.; writing—original draft preparation, X.L.; writing—review and editing, X.L. and Q.L.; visualization, X.L.; supervision, Y.Q.; project administration, Y.Q.; funding acquisition, Y.Q. All authors have read and agreed to the published version of the manuscript.

Funding: National Natural Science Foundation of China (71601073); General Project of Shandong Natural Science Foundation (ZR2021MF113); Shandong University of Science and Technology Virtual Simulation Experiment Teaching Project (XNFZZ2021Y02).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: In this study, the data were mainly obtained from the transportation sector and Gao De Map.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Suganya, S.; Raja, S.C.; Venkatesh, P. Simultaneous coordination of distinct plug-in Hybrid Electric Vehicle charging stations: A modified Particle Swarm Optimization approach. *Energy* **2017**, *138*, 92–102. [\[CrossRef\]](#)
2. Apostolaki-Iosifidou, E.; Codani, P.; Kempton, W. Measurement of power loss during electric vehicle charging and discharging. *Energy* **2017**, *127*, 730–742. [\[CrossRef\]](#)
3. Franke, T.; Krems, J.F. Understanding charging behaviour of electric vehicle users. *Transp. Res. Part F Traffic Psychol. Behav.* **2013**, *21*, 75–89. [\[CrossRef\]](#)
4. Xu, M.; Meng, Q.; Mannering, F. Optimal deployment of charging stations considering path deviation and nonlinear elastic demand. *Transp. Res. Part B Methodol.* **2020**, *135*, 120–142. [\[CrossRef\]](#)
5. He, J.; Yang, H.; Tang, T.Q.; Huang, H.J. An optimal charging station location model with the consideration of electric vehicle's driving range. *Transp. Res. Part C Emerg. Technol.* **2018**, *86*, 641–654. [\[CrossRef\]](#)
6. Xu, M.; Yang, H.; Wang, S.A. Mitigate the range anxiety: Siting battery charging stations for electric vehicle drivers. *Transp. Res. Part C Emerg. Technol.* **2020**, *114*, 164–188. [\[CrossRef\]](#)
7. Li, S.Y.; Huang, Y.X.; Mason, S.J. A multi-period optimization model for the deployment of public electric vehicle charging stations on network. *Transp. Res. Part C* **2016**, *65*, 128–143. [\[CrossRef\]](#)
8. Riemann, R.; Wang, D.Z.W.; Busch, F. Optimal location of wireless charging facilities for electric vehicles: Flow-capturing location model with stochastic user equilibrium. *Transp. Res. Part C* **2015**, *58*, 1–12. [\[CrossRef\]](#)
9. Pan, L.; Yao, E.J.; Yang, Y.; Zhang, R. A location model for electric vehicle (EV) public charging stations based on drivers' existing activities. *Sustain. Cities Soc.* **2020**, *59*, 102192. [\[CrossRef\]](#)
10. Chen, T.D.; Kockelman, K.M.; Khan, M. Locating electric vehicle charging stations: Parking-based assignment method for Seattle, Washington. *Transp. Res. Rec.* **2013**, *2385*, 28–36. [\[CrossRef\]](#)
11. Liu, Q.; Liu, J.H.; Le, W.W.; Guo, Z.X.; He, Z.G. Data-driven intelligent location of public charging stations for electric vehicles. *J. Clean. Prod.* **2019**, *232*, 531–541. [\[CrossRef\]](#)
12. Yang, J.; Dong, J.; Hu, L. A data-driven optimization-based approach for siting and sizing of electric taxi charging stations. *Transp. Res. Part C* **2017**, *77*, 462–477. [\[CrossRef\]](#)
13. Shi, R.F.; Liu, J.H.; Liao, Z.H.; Niu, L.; Ibrahim, E.; Fu, F. An electric taxi charging station planning scheme based on an improved destination choice method. *Energies* **2019**, *12*, 3781. [\[CrossRef\]](#)
14. Lin, Y.P.; Zhang, K.; Shen, Z.J.M.; Ye, B.; Miao, L.X. Multistage large-scale charging station planning for electric buses considering transportation network and power grid. *Transp. Res. Part C Emerg. Technol.* **2019**, *107*, 423–443. [\[CrossRef\]](#)
15. Tao, Y.; Huang, M.H.; Yang, L. Data-driven optimized layout of battery electric vehicle charging infrastructure. *Energy* **2018**, *150*, 735–744. [\[CrossRef\]](#)
16. Guo, J. Research on Euclidean Distance Weighted Coefficient of K-means Clustering Algorithm. *Netw. Secur. Technol. Appl.* **2016**, *10*, 74–75.
17. Reddy, C.K.; Vinzamuri, B. A Survey of Partitional and Hierarchical Clustering Algorithms. In *Data Clustering Algorithms and Applications*; Routledge: London, UK, 2013; Volume 87.
18. Sun, Z.Y.; Hu, Y.J.; Li, W.; Feng, S.W.; Pei, L.L. Prediction model for short-term traffic flow based on a K-means-gated recurrent unit combination. *IET Intell. Transp. Syst.* **2022**, *16*, 675–690. [\[CrossRef\]](#)
19. Yazdani, M.; Chatterjee, P.; Pamucar, D.; Chakraborty, S. Development of an integrated decision making model for location selection of logistics centers in the Spanish autonomous communities. *Expert Syst. Appl.* **2020**, *148*, 113208. [\[CrossRef\]](#)
20. Liu, J.L.; Ong, G.P.; Chen, X.Q. GraphSAGE-based traffic speed forecasting for segment network with sparse data. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 1755–1766. [\[CrossRef\]](#)
21. Kim, G.S.; Chun, J.; Kim, Y.; Kim, C. Coastal Tourism Spatial Planning at the Regional Unit: Identifying Coastal Tourism Hotspots Based on Social Media Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 167. [\[CrossRef\]](#)
22. Zhang, S.; Wang, H.; Zhang, Y.F.; Li, Y.Z. A novel two-stage location model of charging station considering dynamic distribution of electric taxis. *Sustain. Cities Soc.* **2019**, *51*, 101752. [\[CrossRef\]](#)
23. Straka, M.; Buzna, L. Clustering algorithms applied to usage related segments of electric vehicle charging stations. *Transp. Res. Procedia* **2019**, *40*, 1576–1582. [\[CrossRef\]](#)
24. Liu, J.Y.; Peper, J.; Lin, G.; Zhou, Y.; Awasthi, S.; Li, Y.; Rehtanz, C. A planning strategy considering multiple factors for electric vehicle charging stations along German motorways. *Int. J. Electr. Power Energy Syst.* **2021**, *124*, 106379. [\[CrossRef\]](#)
25. Gilanifar, M.; Parvania, M. Clustered multi-node learning of electric vehicle charging flexibility. *Appl. Energy* **2021**, *282*, 116125. [\[CrossRef\]](#)
26. Zhang, Y.H.; Xu, J.; Li, Q.Y.; Zhou, J.L.; Wang, L.L.; Zhu, Z.H.; Li, Y.; Wang, S.R. The Location and Capacity Method of Electric Vehicle Charging Station Based on Density Peak Clustering. *Power Syst. Prot. Control.* **2021**, *49*, 132–139.
27. Sánchez, D.G.; Tabares, A.; Faria, L.T.; Rivera, J.C.; Franco, J.F. A Clustering Approach for the Optimal Siting of Recharging Stations in the Electric Vehicle Routing Problem with Time Windows. *Energies* **2022**, *15*, 2372. [\[CrossRef\]](#)