ISSN 2616-5775 Vol. 8, Issue 1: 64-70, DOI: 10.25236/AJCIS.2025.080109

Spatial Coverage Target-Oriented Matching Algorithm for Ride-Hailing Research

Yitong Zhang[1,a,*]

[1]School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, Sichuan, China
[a]ytz136631@163.com
[*]Corresponding author

Abstract: *As a new type of transportation, ride-hailing not only promotes the development of urban transportation, but also raises some new problems for urban management and management. When ride-hailing are empty, they often go to areas with high order rates to seek maximum benefits, so that they cannot meet the travel requirements of passengers under certain conditions. Due to the lack of a reasonable incentive mechanism, drivers often take negative measures against the "dispatch orders" required by the platform due to refusal to accept orders, shutdowns, and other reasons. In response to this problem, this paper starts with improving the coverage of ride-hailing, and induces the operating paths of ride-hailing without one-way trips by regulating the reward amount of each traffic area, so as to reduce their service in urban operations. Difference between allocation and target passenger distribution. This paper uses the concept of bipartite graph and greedy selection to determine the location of ride-hailing, and then achieves the purpose of improving the coverage of ride-hailing. Finally, through experiments, the model is verified.*

Keywords: *Ride-hailing, Matching Algorithm, Bipartite graph*

1. Introduction

As an emerging, personalized and customized travel mode, ride-hailing has brought great convenience to the public's travel. By 2024, China's online taxi market has initially formed a market structure with Didi Chuxing as the leader, Caocao, Meituan Dache and other major competitors. According to market research data, the number of online taxi users has exceeded 400 million. In the development of ride-hailing services, the issue of spatial coverage has led to frequent difficulties in hailing a ride. First, in popular locations like airports and train stations, passengers often experience difficulty hailing a ride due to traffic congestion or a shortage of available vehicles. Second, in remote areas with fewer orders, drivers tend to reject requests because of longer idle times and lower earnings. Lastly, ride-hailing difficulties occur during both peak and off-peak hours.

In recent years, various methods have been developed to improve task allocation efficiency and worker coverage in spatial crowdsourcing scenarios.Tong[1] et al. a new Global Online Micro-task Allocation (GOMA) problem, develops a two-phase framework with scalable algorithms (TGOA-Greedy and TGOA-OP), and demonstrates through experiments that TGOA-OP outperforms state-of-the-art methods in both efficiency and stability. Similarly, Zhang[2] et al. combined crowdsensing and mobile crowdsourcing to propose a spatial task allocation framework, SpatialRecruiter, using two functions to estimate workers' coverage potential. They designed a heuristic task allocation method, which outperformed traditional spatial task allocation methods in experiments. Zhu[3] et al. investigated the role of vehicle mobility in base station deployment, proposing a greedy-based algorithm that significantly improved delay-constrained coverage in urban areas. Song[4] et al. addressed the issue of inconsistent task coverage by proposing the cTaskMat model, which optimized task coverage and utilized worker preferences to ensure task assignment. Additionally, Wu[5] et al. developed a crowdsensing-based task allocation framework for sweeping coverage, introducing participant incentive models and the PTOS and BGPS task allocation algorithms to maximize social welfare. Ai[6] et al.proposed a multi-step prediction model (PSA-DM) based on historical ride-hailing data, utilizing an encoder-decoder framework and graph attention networks to optimize vehicle scheduling by accurately forecasting ride-hailing demand.

Although ride-hailing services are widely used, research specifically addressing the issue of spatial coverage remains relatively scarce. To tackle this problem, some studies have proposed optimization schemes for resource allocation and vehicle dispatching. However, these approaches still face limitations

ISSN 2616-5775 Vol. 8, Issue 1: 64-70, DOI: 10.25236/AJCIS.2025.080109

in practical applications. Therefore, this paper presents a driver selection model aimed at optimizing spatial coverage for ride-hailing services. By implementing efficient resource allocation and driver dispatch strategies, the model significantly enhances vehicle coverage in different areas and alleviates the issue of ride-hailing difficulties.

2. System Model and Problem Formulation

2.1 Problem Statement

When passengers submit their travel and preference information via internet-enabled devices such as smartphones, including details like travel time, destination, and whether they agree to carpooling, the platform first filters potential drivers based on the straight-line distance from vehicle location points to the order location point. This filtering process is conducted with reference to the traffic zones that the vehicle location points can serve. At this stage, a single vehicle location point may correspond to multiple traffic zones. To handle this, a "bipartite graph" selection method is applied, ensuring that each vehicle location point serves only one traffic zone, avoiding overlaps in assigned zones.

Once this step is completed, a fixed correspondence between vehicle location points and traffic zones is established, where each vehicle location point exclusively serves one traffic zone, and there is no overlap in available vehicle points for any traffic zone. Therefore, to maximize coverage, each traffic zone's spatial coverage problem can be addressed individually. After the initial filtering and allocation, a set of available vehicle location points is formed for each traffic zone. This set can represent the maximum coverage or multiple potential maximum coverage sets. Using a greedy selection algorithm, the fewest vehicle location points are chosen while still satisfying the requirement of maximizing coverage or the maximum coverage set for that traffic zone.

2.2 Model Analysis

2.2.1 Optimization Objective

The optimization objective for maximizing spatial coverage in ride-hailing services mainly includes the following two aspects:

1) Establishing the correspondence between vehicle location points and order locations: This involves ensuring a one-to-one correspondence, meaning that each vehicle location point serves only one traffic zone, and no overlap occurs in the available vehicle location points for any given traffic zone.

2) Maximizing vehicle spatial coverage: After determining the correspondence between vehicle location points and order locations, the goal is to maximize spatial coverage with the selected vehicles.

2.2.2 Constraints

The constraints for maximizing spatial coverage in ride-hailing services are as follows:

1) Vehicle location points and traffic zones: A vehicle location point can serve multiple traffic zones, but it can only be assigned to one traffic zone at a time, while a traffic zone can be assigned multiple vehicle location points.

2) Maximizing spatial coverage with the minimum number of vehicle location points: Once the correspondence between vehicle location points and traffic zones is established, the objective is to select the minimum number of location points that satisfy the maximum spatial coverage.

2.2.3 Assumptions

The assumptions for maximizing spatial coverage in ride-hailing services are as follows:

1) All order locations are assumed to be at the centroid or representative point of the traffic zone.

2) All ride-hailing vehicles are available during the task period.

3) The value of all ride-hailing vehicles to the platform is assumed to be the same within this time period.

4) Ride-hailing drivers prefer to take the shortest path to reach the order location.

5) The distance between a vehicle location point and the order or traffic zone centroid is the straight-line distance between the two points.

ISSN 2616-5775 Vol. 8, Issue 1: 64-70, DOI: 10.25236/AJCIS.2025.080109

2.3 Model Solution

2.3.1 Bipartite Graph for Vehicle-Location Correspondence

Zhang[2] et al. constructed a bipartite graph to represent the relationship between workers and query tasks, considering the task preferences of the workers. A similar situation exists between ride-hailing vehicles and orders. Each traffic zone has a representative point, and the vehicle location points and the representative points of traffic zones can also form a bipartite graph to represent the correspondence between the two.

In real-life situations, vehicles tend to gather at certain location points. Thus, a bipartite graph can be used to preliminarily identify the set of available vehicle locations for each traffic zone. Traffic zones are divided based on demand, with priority given to zones based on the number of edges they connect to in the bipartite graph. The fewer edges a traffic zone has, the higher its priority. The relationship between vehicle location points and traffic zones is such that one vehicle location point can serve multiple traffic zones but can only be assigned to one at a time, while one traffic zone can be assigned multiple vehicle location points. Therefore, a bipartite graph can be constructed to represent the overall structure between vehicle location points and traffic zones.

Specifically, let $L = l_1, l_2, \ldots, l_n$ and $Z = z_1, z_2, \ldots, z_m$ represent the sets of available vehicle location points and traffic zones, respectively. The usable vehicle location points for a traffic zone are determined by multiplying the average waiting time of passengers in traffic zone ii by the average ride-hailing speed in that zone, as shown in Figure 1. Each vehicle location point and traffic zone form a node in the bipartite graph, and if a vehicle location point $l \in L$ can be used by a traffic zone $z \in Z$, an edge $e = (l, z)$ is established between them. These edges form an edge set EE, which leads to a bipartite graph $G = (L, Z, E)$ that describes the relationships between ride-hailing vehicles, traffic zones, and the edge set. The traffic zones are sorted in ascending order based on the number of edges they have, as zones with fewer connections have fewer vehicle options and therefore higher priority for processing.

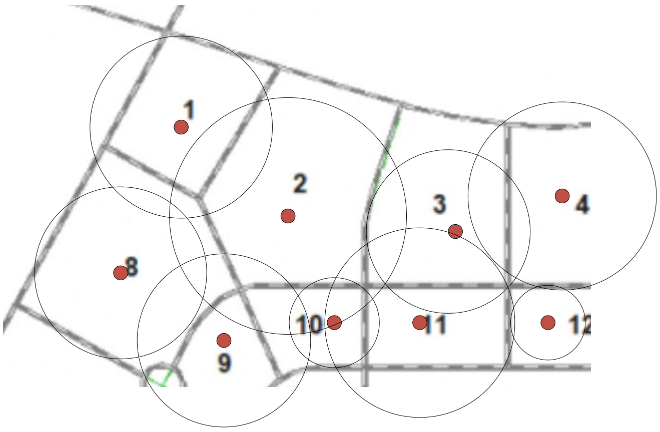

Figure 1: Illustration of Vehicle Location Point Division.

2.3.2 Angle-Based Estimation Function

In the angle-based estimation function, it is assumed that scattered order location points are gathered at a representative point, and each vehicle location point has a potential coverage angle ϕ. These angles can form a maximum coverage area or a maximum coverage area set. The functional form can be expressed as:

$$C = \cup C_j \tag{1}$$

To model the coverage estimation function, we assume that ride-hailing drivers prefer the shortest path to the order location. Although direct routes may not always be available, drivers can choose routes that are closest to the shortest path. Based on this assumption, the shortest path distribution can be represented as a sector radiating from the order location to the vehicle location point. The farther the vehicle location point is from the order, the more route options it has, while closer points have fewer options.

In real-life scenarios, if a vehicle location point l is assigned to an order location O, the sector covering l and O represents the vehicle's coverage potential. The potential angle ϕ reflects the diversity of $l's$ paths. A larger ϕ indicates that the vehicle is farther from the order and may have more

ISSN 2616-5775 Vol. 8, Issue 1: 64-70, DOI: 10.25236/AJCIS.2025.080109

diverse route options, thus covering a larger area.

Based on this analysis, we can design an angle-based coverage estimation function for the candidate vehicle location points. Given an order location o and a set of vehicle location points $L = \{l_1, l_2, \dots, l_n\}$, where n is the number of vehicle location points, if the sectors of selected vehicle location points overlap significantly, their actual trajectories are likely to overlap as well. Therefore, we can choose k vehicle location points that maximize the angular interval around order location o. Using a greedy selection algorithm, the smallest set of vehicle location points can be chosen to satisfy the maximum coverage, ensuring that the selected vehicles have the highest probability of covering the maximum possible area around the target region. For example, as shown in Figure 2, it can be observed that the sectors of l_1、l_2 and l_3 partially overlap, and the total area of the sectors l_1, l_2, l_3 is equal to the total area of l_1, l_2. Since the total area of l_1, l_2 already satisfies the maximum coverage area that can be covered by the three vehicle location points, selecting the ride-hailing vehicles at location points l_1, l_2 is more likely to achieve maximum coverage.

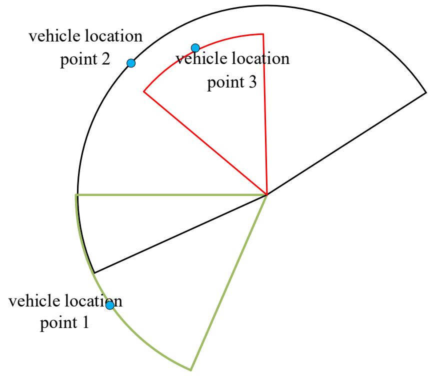

Figure 2: Illustration of Angle-Based Estimation Function.

2.3.3 Greedy Selection

The set of available vehicle location points for a traffic zone can form a maximum coverage or multiple potential maximum coverage sets. A greedy strategy can be used to iteratively select the minimum number of ride-hailing vehicle location points to maximize coverage. The algorithm steps are as follows:

1) Insert all ride-hailing vehicle location points into the set, and calculate the angular coverage of all points.

2) Sort the vehicle location points by distance, from farthest to closest. Form a new set by selecting points from closest to farthest, and recalculate the angular coverage of the remaining points. If the angular coverage remains the same, remove the selected point from the set. If not, return the point to the set.

3) Continue the selection process until the new set $\overline{L}_i$ is obtained.

If the demand in traffic zone ii exceeds the available number of vehicles, no further processing is required.

Given:

- $L = \{l_1, l_2, \dots, l_n\}$ is the set of n vehicle location points.
- $\phi(l_i)$ represents the angular coverage of vehicle location point l_i.

The goal is to find a subset $\overline{L} \subseteq L$ such that the total angular coverage $\Phi(\overline{L})$ is maximized while minimizing the number of points in $\overline{L}$.

The pseudocode of the greedy algorithm is as follows:

1) Initialize an empty set:

$$\overline{L} = \emptyset \tag{2}$$

2) While the angular coverage is not maximized:

Select the vehicle location point $l_i \in L$ that provides the maximum marginal increase in angular coverage:

ISSN 2616-5775 Vol. 8, Issue 1: 64-70, DOI: 10.25236/AJCIS.2025.080109

$$l_i = \arg\max_{l \in L \setminus \overline{L}} \left(\Phi(\overline{L} \cup \{l\}) - \Phi(\overline{L}) \right) \tag{3}$$

Update the set $\overline{L}$:

$$\overline{L} = \overline{L} \cup \{l_i\} \tag{4}$$

3) Stop when the maximum coverage is achieved or no further improvement can be made.

The total angular coverage for the selected set $\overline{L}$ is:

$$\Phi(\overline{L}) = \sum_{l \in \overline{L}} \phi(l) \quad \text{subject to} \quad \text{no significant overlap in } \phi(l). \tag{5}$$

This algorithm continues until the angular coverage cannot be further increased, ensuring that the fewest number of vehicle location points achieve maximum coverage.

3. Numerical Study

3.1 Simulation Setup

Select the map of the desired research area from OpenStreetMap and divide the selected area into regions. The endpoints of the selected area are labeled, as shown in Figure 3. Define the area enclosed by labels A, B, C, D, and E as Region 1, and the area enclosed by labels C, D, E, F, G, H, I, J, and K as Region 2. Random ride-hailing vehicle location points are generated in each region, and it is assumed in the experiment that each ride-hailing vehicle location point covers the same angle, with the reference direction set to due north.

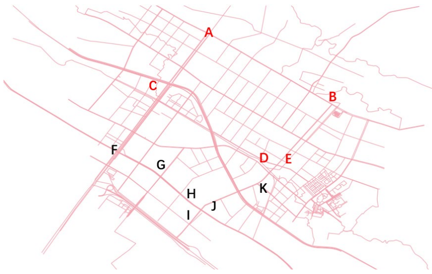

Figure 3: Illustration of Angle-Based Estimation Function.

3.2 A Vehicle Location Allocation Method Based on Bipartite Graphs

We utilized the GeoGraphiclib library in Python to accurately calculate the distance between vehicle location points and representative points of traffic zones. Based on the scale of the study area, a service radius of 2 kilometers was selected for the vehicle location points. The edges of traffic zones were defined as the product of the average waiting time of the zone and the average speed of ride-hailing taxis.

By calculating the relationship between available vehicle location points and traffic zones, we employed a greedy selection algorithm to allocate vehicles to each zone. The algorithm prioritizes zones with higher priority, ensuring that their limited available vehicle resources are allocated first. Table 1 shows the vehicle location point allocation results for Region 1 and Region 2.

Table 1: Vehicle location point allocation results for Region 1 and Region 2.

Region No.	Available Vehicle Location Point IDs
Region 1	[4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]
Region 2	[1, 2, 3, 17, 18, 19, 20, 21, 22, 23, 24]

For cases where location points overlap (e.g., vehicle location points 1, 2, and 3 appear in both Region 1 and Region 2), we further optimized the vehicle allocation scheme to ensure maximum vehicle utilization efficiency across the regions.

3.3 Vehicle Location Point Coverage Angle ϕ Set to 30°

Using a greedy algorithm, we can obtain the maximum coverage range of ride-hailing vehicles and the allocation results when the coverage angle ϕ is 30°. Table 2 provides the maximum coverage range,

ISSN 2616-5775 Vol. 8, Issue 1: 64-70, DOI: 10.25236/AJCIS.2025.080109

removable vehicle location points, and final confirmed vehicle location points for each region, as shown in below.

Table 2: Allocation results of ride-hailing vehicle location points.

Region No.	Max Coverage	Removable Vehicle Location Points	Final Points
Region 1	[(130.5758, 286.8271), (288.9648, 318.9648)]	14, 16, 10, 8	6, 4, 9, 7, 15, 13, 12, 11
Region 2	[(21.8307, 51.8307), (97.6719, 127.6719), (128.0899, 158.0899), (195.0909, 300.8475)]	2, 19	1, 3, 17, 18, 20, 21, 22, 23, 24

3.4 Analysis of the Impact of ϕ Value

For Region 1, a comparative experiment was conducted with the coverage angle ϕ set at 15°, 30°, 45°, and 60°, respectively, and the results are presented in Table 3 below.

Table 3: Comparisons of different angles in Region 1.

Angle (°)	Max Coverage	Removable Vehicle Location Points	Final Points
15	[(29.3306626, 44.33066257), (105.1719338, 120.1719338), (135.589897, 150.589897), (165.933552, 180.9335519), (202.590982, 243.1744086), (245.762976, 260.7629759), (270.799564, 293.3474687)]	[]	[1, 2, 3, 17, 18, 19, 20, 21, 22, 23, 24]
30	[(21.8307, 51.8307), (97.6719, 127.6719), (128.0899, 158.0899), (195.0909, 300.8475)]	2, 19	[1, 2, 3, 17, 18, 19, 20, 21, 22, 23, 24]
45	[(14.3306626, 59.33066257), (90.1719338, 308.3474687)]	[19, 2, 20, 21]	[17, 3, 1, 18, 22, 24, 23]
60	[(6.8306626, 66.83066257), (82.6719338, 315.8474687)]	[19, 20, 21, 2]	[17, 3, 1, 18, 22, 24, 23]

As the coverage angle ϕ increases, more ride-hailing vehicle location points can be filtered out, resulting in a larger coverage area. When ϕ is set to 30°, the distribution of the ride-hailing vehicle location points is more dispersed. However, the maximum coverage area does not significantly differ between ϕ set at 45° and 60°.

4. Conclusion

In this paper, we presented a spatial coverage target-oriented matching algorithm for ride-hailing services, focusing on improving the efficiency and fairness of vehicle allocation across different traffic zones. The proposed approach utilizes a bipartite graph structure to establish a one-to-one correspondence between vehicle location points and traffic zones, while a greedy selection algorithm ensures that the minimum number of vehicles are used to maximize coverage. Through the use of an angle-based estimation function, we were able to evaluate and optimize vehicle positioning for ride-hailing services, ensuring better coverage even in areas with lower demand. Our experiments demonstrated that the proposed method significantly improves ride-hailing service coverage compared to traditional allocation methods, providing a practical solution for urban transportation challenges related to ride-hailing inefficiencies. This research contributes to the optimization of ride-hailing platforms, reducing the imbalance in service availability across urban and remote areas, and ultimately improving the overall user experience. Future research may explore dynamic adjustments to the model to account for real-time changes in demand and traffic conditions.

ISSN 2616-5775 Vol. 8, Issue 1: 64-70, DOI: 10.25236/AJCIS.2025.080109

References

[1] Tong Y, Zeng Y, Ding B, et al. Two-sided Online Micro-Task Assignment in Spatial Crowdsourcing [J]. IEEE Transactions on Knowledge and Data Engineering, 2020: 1-1.
[2] Zhang X, Yang Z, Gong Y-J, et al. SpatialRecruiter: Maximizing Sensing Coverage in Selecting Workers for Spatial Crowdsourcing [J]. IEEE Transactions on Vehicular Technology, 2017, 66(6): 5229-5240
[3] Zhu Y, Bao Y, Li B. On Maximizing Delay-Constrained Coverage of Urban Vehicular Networks [J]. IEEE Journal on Selected Areas in Communications, 2012, 30(4): 804-817
[4] Song S, Liu Z, Li Z, et al. Coverage-Oriented Task Assignment for Mobile Crowdsensing [J]. IEEE Internet of Things Journal, 2020, 7(8): 7407-7418
[5] Wu L, Xiong Y, Wu M, et al. A Task Assignment Method for Sweep Coverage Optimization Based on Crowdsensing [J]. IEEE Internet of Things Journal, 2019, 6(6): 10686-10699
[6] Ai. Efficient Multi-Step Prediction Model That Considers the Influence of Spatial and Temporal Factors on Ride-Hailing Demand [J]. TRANSPORTATION RESEARCH RECORD, 2024.