2016 Joint International Conference on Artificial Intelligence and Computer Engineering (AICE 2016) and International Conference on Network and Communication Security (NCS 2016)
ISBN: 978-1-60595-362-5

Relay Carpool Method Based on Location Data Matching

Chun LIU[a,*], Wu LIANG, Meng-Xi TAN

Institution of Computer Science and Technology,
Hubei University of Technology, Wuhan, Hubei, China

[a]251207905@qq.com

[*]Corresponding author

Keywords: Data Mining, Clustering, Relay-carpooling, Location Service, Matching Degree.

Abstract. The reasonable carpool algorithm can alleviate the traffic congestion greatly, optimize passenger's experience, and reduce carbon emissions. In this paper, we propose a two-stage relay carpool algorithm. In the first phase, the carpooling requirements are assigned to each vehicle based on passenger assignation algorithm of improved Hausdorff distance, so the multiple vehicles demand problem is transferred to the single vehicle requirement question. Then in the second phase, based matching degree the clustering results chooses the best matching vehicle as the candidate. The experiments show that this method has achieved the goal to make recommendations for both single vehicle carpooling and multi- vehicle carpooling.

Introduction

Carpool is the phenomenon that people who reach the same routine or destination by way of sharing vehicles [1], which has capability of solving the difficult problems of alleviating the traffic congestion greatly, reducing the quantity of the vehicle flow, and effectively using of resources for vehicles. In this paper a carpooling match recommendation scheme is proposed without modifying original routine. Based the location data uploading from request client, firstly we extract direction changing waypoint to simplify the process of multi-routines, then assign carpooling passengers to different vehicles based improved Hausdorff distance, after that a clustering process based match degree is performed to choose the best matching vehicle as the optimum recommendation candidate. The experiments show that this method and process can give the optimum recommendations for both single-vehicle carpooling and multi-vehicle relay carpooling with the constraints of time windows and original routine, achieve the bidirectional preferable between vehicles and carpooling passengers.

Commute Carpooling Model

Suppose that there are some vehicles which provide carpooling service in an area, at the same time multi-vehicle relay carpooling demands appears, but each has the waiting time constraint. Each vehicle can accept the maximum demand amount no more than their maximum number of passengers, which means that we do not consider the vehicle seat reuse condition; besides passengers can walk to a pickup point in the range of walking distance. In this paper our question is defined to real-time dynamic recommend the carpooling scheme based original routine and meeting the time window constraint, and our target of solving this problem is achieving the highest success rate of carpool.

Vehicle routing is a sequence of location data organized by time series [2], which can be viewed as a sequence (x,y,t) including time and position information. Therefore, the definition of travel routes of private vehicles in the area is $TR_i = \{(x_k,y_k,t_k)\}$, here stands for the identification of private car, $1 \le i \le n$; k means the No. k location record, while the location information can also be expressed as v= (x, y). The set of traffic route is TR_n.

Assumed that the carpooling demands in this region at the moment is $TR_j=(s_j,d_j,t_j)$, where j is the serial number of demand, $1 \leqslant j \leq m$. s_j,d_j is respectively the location coordinates corresponding to the starting point and the end point (also known as the source point s, destination point d), t_j represents the car arriving or pick-up time appointed by the demand. If the distance between passenger source location and get-in location is within walking distance, we call the pick-up location is in the neighborhood field of the source point (represents as S). Similarly the destination neighborhood field D can be defined.

The carpool can be divided into single time carpool and multiple trips interchanging carpool based on relay mechanism according to the passenger can nonstop arrive the target or not. The goal of this paper is to obtain the highest carpooling success ratio and the lowest total cost of the vehicle on the basis of the un-limitation of interchanging condition. The problem can be formulated as two objective functions: $\max \sum_{i\in C}\sum_{x\in v}\sum_{y\in v} X^{i}_{x,y}$ and $\min \sum_{j\in C} \mathrm{Cos}\,t_j$ Here (1) $X^{i}_{x,y}$ is a binary variable, v is a ride location set where the passenger can take the car on the vehicle trajectory. C is on behalf of the vehicles set. X $X^{i}_{x,y}=1$ means the vehicle i can travel from vertex x to vertex y (with direction selective). The minus value means the lowest cost of demands for the passenger j.

Carpooling model can be divided into single time carpool and relay transferring carpool. If taking the ride location as the center, both the demand location neighborhood field and target location neighborhood field within passengers walking distance as the radius have intersections with track traffic, then this demand can be simplified as a non-transferring single time carpool mode. In the carpool illustrative figure 1, the circles represent the location point on the way, here the passenger P1 and vehicle C1 have the same source point neighborhood field and destination point neighborhood field; passenger P2 and vehicle C1 have the same destination point neighborhood field, while the source point neighborhood field and the trajectories of vehicle C1 have intersect at the point 2; the source point neighborhood field and the destination point neighborhood field of passenger P3 have intersects at the point 4, 6 respectively with the trajectories of vehicle C1, and the source point neighborhood field. The upper descripted situations can all be realized carpooling demands through the single time carpooling process.

a) Non-Transferring single time carpool scheme b) Multiple time carpool scheme based transferring

Figure 1. Single time and Multiple time carpool scheme.

If the source point neighborhood field and the destination point neighborhood field of carpooling demands have intersections with trajectories of different vehicle owners, also different trajectories have intersections, then car passenger can ride and transfer different vehicles at the intersection location to reach the destination, which belongs to the transferring needed multi-trips carpool. In figure 1(b), if the trajectories of C1 and C2 are respectively (1-2-3-4-5), (6-7-3-8), the carpooling demand P4 can take C2 car firstly, then transfer C1 car at point 3 to reach to the destination.

Relay Carpool Schedule

For the single time carpool and multiple trips transferring carpool problem, in this paper the solution is shown as follows: Firstly, extract the turning point of the travel trajectory as the simplified description of the trajectory; Then according to the passenger matching algorithm based on Hausdorff distance, assign the carpooling demands to the corresponding vehicles; Finally, introducing the transferring points, further select the most matching demands through clustering based on the matching degree, and push the carpool schedule to the passengers according to their cost order.

Vehicle Route Pretreatment

The processing of trajectory has heavy computation and complex calculation process. Therefore we extract the turning points to descript trajectories. These points are the positions at which vehicles change the direction, at the same time, the direction angle in the GPS records have large changes [3]. In this paper we sort the vehicle trajectory data by ascending time, and denote the point of trajectory by v_i. So the angle difference of the adjacent two points by time order, v_i and v_{i+1}, can be computed as:

$$\Delta\alpha = \left| GPSDerection_i - GPSDerection_{i+1} \right| . \tag{1}$$

If $70^o < \Delta\alpha < 110^o$, $250^o < \Delta\alpha < 290^o$, then we judge the point according to the angles as a turning point. After the extraction and reconstruction the trajectory with turning points, the driving trace of vehicles can be compressed into a sequence $\{(v_i,t_i)\}$ with timestamp, which makes the amount of trace data greatly reduced.

Passenger Matching Algorithm Based on Improved Hausdorff Distance

An improved MHD (Hausdorff Distance Modified) algorithm based on Hausdorff distance [4] can be used to calculate the matching degree among the points set. Although the turning sequence set of the vehicle trajectory belongs to a set of points, it has obvious characteristics of the line segment. So the shortest distance from points in set A to line segments in set B have better characteristics to reflect the position of the matching degree than that matching degree of point set obtained with the MHD algorithms. Therefore, we improve the MHD, and define $h(a,TR_i)$ as the shortest distance from the point a to the line segment in set TR_i. Given the points set TR_j and trace line segment set TR_i, the improved Modified Hausdorff Distance, h_{imhd} is[5]:

$$h_{imhd}(TR_j, TR_i) = \frac{1}{N_a} \sum_{a \in TR_j} h\left(a, TR_i\right) . \tag{2}$$

Here Na is the account of points in TR_j, $h(a, TR_i)$is the shortest distance from point a to line segments in TR_i. While the carpooling demand in our research can be seen as a special trajectory TR_j which only includes source location point V_s and destination point V_d, so the constraint formula of carpooling trace is:

$$h_{imhd}(TR_j, TR_i) = \frac{1}{2} h\left(v_s, TR_i\right) + \frac{1}{2} h\left(v_d, TR_i\right) \le s . \tag{3}$$

The formula shows that if the distance from the carpooling demand location to pick-up location or get-out location of the same trajectory is within the acceptable range of walking distances, then the single vehicle carpooling demand can be fulfilled. When $h_{imhd}(TR_j, TR_i) > s$, if the distances between the carpooling demand location and the getting in location of trajectory TR_i and TR_{i+1} are all within the acceptable range of walking distance, and there is intersect point between these two trajectories, then the demand can be fulfilled by multiple interchanging. So the constraint formula of multiple interchanging carpools is:

$$h_{imhd}\left(TR_j, \left(TR_i \cap TR_{i+1} \ldots \cap TR_{i+k-1}\right)\right) = \frac{1}{2} h\left(v_s, TR_i\right) + \frac{1}{2} h\left(v_d, TR_{i+k-1}\right) \le s . \tag{4}$$

Here, $TR_i \cap TR_{i+1} \ldots \cap TR_{i+k-1} \ne \Phi$ means that there exist intersects between the trajectories of k different cars, TR_i, TR_{i+1},…, TR_{i+k-1}.

In the events of commuter carpool, the acceptable waiting time of the passenger P and the vehicle C respectively are limited. Based this limitation, the constraint conditions about the passenger's getting in time stamp and really vehicle arrival time stamp would be given in the time dimension, in another word, the demand and the service have approximation in the time dimension:

$$TSim(v_s, TR_i) = T_{TR_i} - T_{v_s}, \left|\mathrm{TSim}(\mathrm{v_s}, \mathrm{TR_i})\right| \le T . \tag{5}$$

In summary, taking into account the influences of route factors and time factors on the carpooling demand and service, the following definitions are given: Supposed that the longest walking distance s=500m,the maximum waiting time T=10 min, if the trajectory corresponding to passenger's demand $TR_j \in TR_m$, and vehicle service trajectories $\{TR_i \cap TR_{i+1} ... \cap TR_{i+k-1}\} \in TR_n$, their space distances are less than S, and the time differences are less than T,then we call trajectory TR_j as the time-space neighbor of $TR_i, TR_{i+1} ..., TR_{i+k-1}$.

$$N(TR_j) = \left\{ TR_j \in TR_m \middle| \begin{array}{l} TSim(v_s, TR_i) \le T \\ h_{imhd}\left(TR_j, (TR_i \cap TR_{i+1} ... \cap TR_{i+k-1})\right) \le S \end{array} \right\}. \tag{6}$$

$N(TR_j)$ is the carpooling demands set nearby service routes $TR_i, TR_{i+1} ..., TR_{i+k-1}$ whose location can meet spatial accessibility and time effectiveness simultaneously. With the above pretreatment the carpooling demands can be assigned to specific vehicles.

Clustering Algorithm Based on Matching Degree

Although the carpooling demands being assigned to different vehicles, the service capabilities of vehicles still have their limitation. So the service queues and sorting for vehicles are needed, and it is necessary to choose those demands with much more matching degree to vehicles trajectories from queues as the recommendations to both sides of the carpooling transaction. In this paper, based on matching degree we use clustering algorithm to further filter demands, and choose the carpooling demands with better matching as the recommendation candidates.

Proportional selection operator method is a kind of selection algorithm based on the proportion, which decides the possibilities of individuals' descendants survival according to the proportions of each individual's adaptability. Therefore we introduce the concept of matching degree [6] to describe the relationship between the vehicle trajectories and the carpooling demands, and give the following definition:

*Definition*1. In carpool mode, the difference between passenger getting-in time T_{v_s} and vehicle arrival time T_{TR_i} is $\Delta T_{ij} = T_{TR_i} - T_{V_s}$. So we define the time matching degree about demand P_j and vehicle C_i in time dimension as: $T_{ij} = \beta_1 \cdot \frac{1}{|\Delta T_{ij}|}$. Here $1 \le |\Delta T_{ij}| \le 10$, the time unit is minute, β_1 is the weight, T_{ij} value range is (0,1].

Definition2. Route matching degree H_{ij} about demand P_j and vehicle C_i is defined as the distance ration of passenger taking distance H_j and vehicle traveling distance H_i in this trajectory, $H_{ij} = \beta_2 \cdot \frac{H_j}{H_i}$, β_2 is the weight, the value of H_{ij} belongs to (0,1].

Definition3. In summary, we can define the possibility of a successful carpooling demand P_j matching with service vehicle C_i as the comprehensive matching degree MP_{ij}, and stated as:

$$MP_{ij} = \begin{cases} \beta_1 \cdot \dfrac{1}{\Delta T_{ij}} + \beta_2 \cdot \dfrac{H_j}{H_i}, 1 \le |\Delta T| \le 10 \\ 0, else \end{cases}. \tag{7}$$

MP_{ij} is a comprehensive index that integrates the time matching degree and the route matching degree. The greater the value is, which means the shorter the time of the passengers waiting for the bus is and the greater the route matching degree is, the greater the likelihood of success is.

According to the definition of matching degree, in the single time carpool mode, the matching degree of the demand and the single vehicle can be obtained directly with the equation. While in the multiple trips carpooling mode, one demand needs k vehicles to relay support. So the transfer points can be found at route crossing points hidden in vehicle trajectory sequences TR_i, TR_{i+1}..., TR_{i+k+1}. Hence one demand can be separated to k ordered sub-demand sequences: $P_j=P_i+P_{i+1}+\ldots+P_{i+k-1}$, each sub-demand response to a sub-matching degree of a transfer vehicle, means: $MP_j=MP_i+MP_{i+1}+\ldots+MP_{i+k-1}$, then after separated sub-matching degree can be grouped with Probabilistic Clustering algorithm. Besides considered the actual carpooling actions in commuter time, we limit the transfer times in 2 times, so the probabilistic clustering process based matching degree is [7]:

(1) According to all demands to each vehicle, the total matching degree of the vehicle $\sum_{i=1}^{n}\sum_{j=1}^{m}MP_{ij}$ can be calculated. Here i is the identify number of car, j is the demand serial number for the passenger. (2) Finding the relative fitness of each vehicle responding to carpooling demand, that is, the selected probability is: $MP_{ij}/\sum_{i=1}^{n}\sum_{j=1}^{m}MP_{ij}$. (3)Supposed threshold value $t\in(0,1)$, in accordance with the principle of Roulette, select demands which meet $MP_{ij}/\sum_{i=1}^{n}\sum_{j=1}^{m}MP_{ij}>t$, and then find the best passenger demands to recommend to the service vehicle. (4) For the other side, when a passenger facing a variety of carpooling options, the minus total travel cost value computed based on time and space factors can be viewed as carpooling recommendation standard:

$$\min\sum_{j\in P}\mathrm{Cos}t_j=\sum_{i=1}^{k}f(H_j)+\sum_{i=1}^{k}f(\Delta T_{ij})\,. \tag{8}$$

$\sum_{i=1}^{k}f(H_j)$ is the total cost of riding k different vehicles, $\sum_{i=1}^{k}f(\Delta T_{ij})$ is the fine given by passenger or vehicle when the waiting time of each side exceeds the agreed time. The income of car's owner [8] can be expressed as:

$$\sum_{i\in C}\mathrm{Cos}t_i=f\left(c_i\right)+\sum_{j=1}^{j}f\left(c_j\right)\cdot H_v^j\cdot P_{v-1}^j\,. \tag{9}$$

Here $f(c_i)$ is the cost of vehicle i traveling the trajectory, $f(c_j)$ is the addition cost when an extra carpooling passenger get in at the driving process. H_v^j is the distance from point v *to point* v-1, P_{v-1}^j is the passenger numbers getting in at point v-1.

Experimental Results and Analysis

For verifying the carpool algorithm model and its validity, we use random function to generate 10 units traveling data which only contain road turning points, which descripted as <turning point, time> record fields, such as {(3,17),(6:00);(5,17),(6:09);(5,1),(7:10)}; and 30 items carpooling demand data, which descripted as <demand, start time/start location> record fields, such as {1, (6:22), (3,12),(14,10)}, {2,(2,15),(6:40);(5,15), (6:53); (5,12),(7:05);(9,12), (7:20); (11,11),(7:31)}; then simulate the model with MATLAB.

To illustrate, a trajectory map Figure 2 is drawn to expressed the location and demand data. Here line segment represents the vehicle trajectory, + represents location of the getting in and getting out points, j ↑, j↓ and black dot represents the getting in and getting out points responding to the No. j demand.

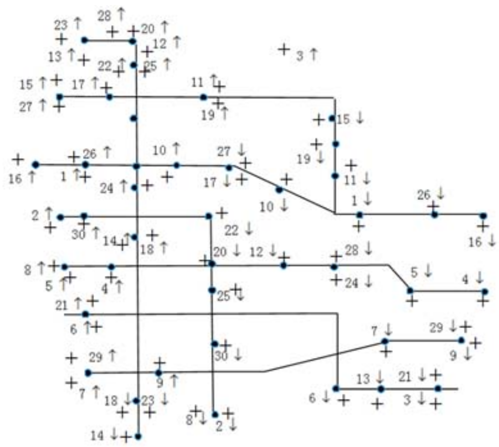

Figure 2. Carpooling trajectory map.

According to the route and the carpooling demand information, the passengers will be assigned to different vehicles based allocation algorithm, and get the first matching relationship of vehicles and carpooling demands, namely, the requirements are assigned into the spatial-temporal neighborhood corresponding to the vehicle trajectory. So a [vehicle, (demands, direction)] matrix can be built as [(1),(23↑,23↓,14↑,14↓,18↑,18↓,13↑,12↑,24↑,28↑,22↑,25↑,20↑)] or [2, (27↑,27↓,17↑,17↓,10↑,10↓)]. Here ↑ and ↓can be represent by index 1 or 0.

Through experimental comparison, when β_1 =0.1and β_2 = 0.9, the match has quite high success rate. After using clustering algorithm, matching degrees of demands and vehicles can be achieved as descripted in table 1.

Table 1. The matching relationship table of demands and vehicles.

vehicle	the first matching degree	the multi-matching degree
1	23/0.85,14/0.5,18/0.333	13/0.683,22/0.313,25/0.311,20/0.364
2	27/0.756,17/0.622,10/0.362	
3	25/0.805,22/0.604,20/0.796	
4	28/0.825,12/0.744,24/0.546,20/0.669	
5	15/0.583,11/0.5,19/0.419	
6	16/0.92,26/0.738,1/0.659	
7	2/0.95, 30/0.743	25/0.398,22/0.205
8	5/0.845,4/0.830	
9	21/0.839,6/0.649	13/0.632,
10	29/0.925,7/0.771,9/0.746	

In Table 1, carpooling demands No. 22 and No. 25 can be accomplished either by single or multiple relay carpools. Generally the single carpool has a higher priority than relay carpooling, namely vehicles provide prior services to the single carpooling demand. Because there is no direct vehicle service for demand No. 13, this demand can only accomplished through the relay carpools with transferring mechanism. As for other requirements, both of them can be completed by single carpool.

According to the matching degree of demands and vehicles in table 1, with the principle of roulette carpooling demands which are satisfied to $MP_{ij} \geq T$ condition will cluster as a class. Here assumed T is 0.5, we can get the carpooling demand set assigned to each vehicle, in another word matching results Table 2. From table 2 the final match relationship between carpooling demand and per car has been obtained eventually, and for passengers, a carpooling plan with the minimum travel cost computed with equation (9) will be chosen to recommend for them. As for the demand No. 18 and No. 10, their matching degrees are lower than the threshold value, which needs to be eliminated from the matching result.

Table 2. Match Result Table.

vehicle	demand	vehicle	demand
1	23,14,13	6	16,26,1
2	27,17	7	2,30
3	25,22,20	8	5,4
4	28,12,24,20	9	21,6,13
5	15,11	10	29,7,9

Experimental results show that this carpool model can not only realize the single time carpool, such as demand 23, 14; but also realize the relay multiple trips carpool, such as demand 13, based transferring mechanism to take the car No. 1, and then transfer to car No.9 at relay point to drive to the destination. Although the above model and mechanism help to realize more flexible carpool, due to the more strict transferring restrictions about vehicle routes and relay times, they reduce the success rate of transferring carpool.

Conclusion

Carpool is a share traveling mode helping to alleviate the current traffic problems by way of sharing idle vehicle resources. In this paper, an improved model of relay multiple trips carpool model is proposed, which introduces concept of turning points to reduce data amount and complexity of trajectory processing, and assign carpool demands to different service vehicles according to passenger assign algorithm, greatly reduce the matching scope; then introduce concepts of transfer point and matching degree to compute how much vehicle routes match to carpooling demands, after that cluster demands centered on vehicles, with the principle of roulette based computed degree values to further optimize range of carpooling demands each vehicle can provide service; finally through the two-way choice between passengers and owners, carpooling scheme suitable to each other can be confirmed. With this relay and recommendation mechanism, whether single time carpool or multiple trips carpool can be very well realized, and meet the requirements from both supply side and demand side.

Acknowledgement

The paper is supported by Natural Science Foundation of Hubei Province China (No: 2014CFB594).

Reference

[1] Cheng Jie, Tang Zhi-hui, Liu Jie. Research on Dynamic taxi pooling model based on Genetic Algorithm [J]. Journal of Wuhan University of Technology. 2013, 1(2): 187-191.

[2] Gordeau J.-F., Laporte G. The dial-a-ride problem: models and algorithms [J]. Ann Oper Res. 2007, 4(1): 29-46.

[3] Wang Zi, Zhou Tong, Wang Yang. Design and implementation of intelligent carpooling system based ongps data visualization [J]. Computer Engineering and Design. 2015, 36(2): 529-533.

[4] Cao Yanyan, Wu Jian.An improved Hausdorff distance and spectral clustering vehicle trajectory pattern learning approcah [J]. Computer Applications and Software. 2012, 5(5): 38-40.

[5] Xia Ying, Liu Shen-yi. Real time trajectory anomaly detection method and its application [J]. Journal of Chongqing University of Posts and Telecommunications. 2011, 4(8): 496-499.

[6] Shao Zeng-zheng, Liu Hong. Multi-Vehicle carpool problem of two stage clustering heuristic optimization algorithm [J]. Computer Research and Development. 2013, 50(11): 2325-2335.

[7] Han J, Kamber M, PEI J. Data Mining:Concepts and Techniques [M]. 3nd ed. China Machine Press. 2012:55-79.

[8] Li Xiang-yong. A Study on Models and Algorithms for Vehicle Routing Problems. The Degree of Doctor of PhilosoPhy ShanghaiJiaotong University. 2007.