Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Procedia - Social and Behavioral Sciences 96 (2013) 1898 – 1909

Procedia Social and Behavioral Sciences

13th COTA International Conference of Transportation Professionals (CICTP 2013)

A Hierarchical Path Planning Method Using the Experience of Taxi Drivers

HU Ji-hua[a]， HUANG Ze[b]， DENG Jun[a,b,] *

[a]*Research Center of Intelligent Transportation System, Sun Yat-sen University, Guangzhou 510006, China*
[b]*Guangdong Provincial Key Laboratory of Intelligent Transportation System, , Guangzhou 510006, China*

Abstract

Effective path planning has been identified as an important requirement for route navigation in Intelligent Transportation Systems (ITS). However, the paths computed by the conventional path planning algorithms are usually not optimal because they ignore the drivers' experience and the characteristics of the local road network. Routes chosen by taxi drivers are believed to be more representative, so it can be significant to take advantage of the taxi drivers' experience to support the path planning. We present a hierarchical path planning method based on the experiential routes of taxis. The algorithm consists of three steps. Firstly, routes are extracted from original taxi trajectories. Secondly, all roads are categorized according to the track data and then the road network is classified into two grades using travel frequency for road segments. Thirdly, combined with Dijkstra algorithm, a hierarchical path planning method which searches paths by traversing the hierarchy is proposed. Guangzhou, China is chosen for the study because of the availability of taxi GPS data. This study compares the paths planned by the proposed approach with the conventional algorithm's results from the aspects such as travel time and route length. The experimental results show that the method proposed incurs much lower costs in travel time than the traditional ITS path planning using Dijkstra or hierarchical Dijkstra algorithms.

Selection and peer-review under responsibility of Chinese Overseas Transportation Association (COTA).

Keywords: Intelligent Transportation Systems; path planning; Dijkstra algorithm; experiential routes of taxi; hierarchical road network

1. Introduction

As one of the key technologies of the Intelligent Transport System（ITS）, path planning algorithms can effectively help the drivers choose the routes, avoid congestion and reduce travel costs. At present, most of the path-planning researches focus on the algorithm efficiency and the shortest path of mathematical meaning, which

* Hu Ji-hua. Tel.: 13662452092.
E-mail address: hujihua@mail.sysu.edu.cn

Selection and peer-review under responsibility of Chinese Overseas Transportation Association (COTA).
doi:10.1016/j.sbspro.2013.08.215

can't truly meet drivers' preference and demands. Paths computed by the conventional path planning algorithms are usually not optimal because they ignore the drivers' experience and do not consider the characteristics of the local road network. As a new type of search strategy, hierarchical search strategy can effectively improve search efficiency and meet the requirements of the driver better. Therefore, the hierarchical search strategy has drawn scholars' close concern.

Hierarchical search strategy is famous in the field of artificial intelligence. Polya first introduced the concept of hierarchy in the road network. Since then, the concept has been further explored by Sacerdoti(1974)、Korf (1987) and many other researchers(Jung S, Pramanik S,2002; Jagadeesh G et al.,2002;Chou Y L, Romeijn H E et al.,1998;Liu B,1997;Car A et al.,1994). Many domestic scholars (WENG M et al.,2006; GAO S et al.,2009; WU X L et al.,2006; ZHONG H L et al.,2011) have further study on it and used hierarchical search strategy to optimize the path planning algorithm. In most of the previous studies, the road network is divided into different classes according to the attribute of the road network; such a manner has two shortcomings: first, the obtained hierarchical road network is static and can't properly reflect the actual dynamic traffic; second, it does not fully consider the expectations of the driver.

In recent years, the importance of the taxi driver's experience is gradually highlighted in the vehicle navigation research at home and abroad. More and more scholars apply it to the dynamic navigation based on the real-time traffic information (TANG L L et al.,2010; YANG Y et al.,2009). Combined with the experience of the taxi driver, path planning algorithm can take advantage of the historical experience to avoid congestion and reduce travel time. The path calculated may not be the shortest path, but it is the optimal path. Therefore, integrating the taxi driver's experience into the path planning algorithm can provide a reliable path.

The remainder of the paper is organized as following. The next section of this paper presents the methods of the study. Section 3 illustrates how to extract the paths of taxi according to the GPS data. In section 4, an experiential hierarchy for the road network is constructed. In section 5, a hierarchical path planning algorithm is proposed based on the experiential road network. In section 6, a case study of Guangzhou City, China is demonstrated by using taxi GPS data. The method proposed is compared with the conventional path planning algorithm from the aspects such as travel time and route length. Section 7 concludes with some brief comments.

2. Methods

The paths chosen by taxi drivers often involve a variety of factors: degree of the congestion, travel time, the surrounding environment and costs. Therefore, this paper combines the taxi driver's experience with the hierarchical search strategy, and optimizes the shortest path algorithm by constructing an experiential road network.

Fig.1 presents a framework for the technical process. The framework consists of three main procedures (rectangular boxes in Fig.1): preprocessing of the floating car data, constructing an experiential road hierarchy based on the experience of taxi drivers, and computation of the experiential optimal path.

The main technical route is as follows:

First, the collected floating car data must be preprocessed. Map-matching technology is used to assign original GPS coordinates to corresponding road segments. Then the taxi routes based on GPS data are extracted.

Second, experiential paths are selected from a large number of taxi driving trajectories by setting a speed threshold. Then calculate the travel frequency of each road segment within a limited time and construct an experiential hierarchy of road network according to the number of times each road segment is traveled.

Finally, on the basis of the experiential hierarchy of road network, a hierarchical path planning algorithm is implemented.

We describe these three procedures in the following sections.

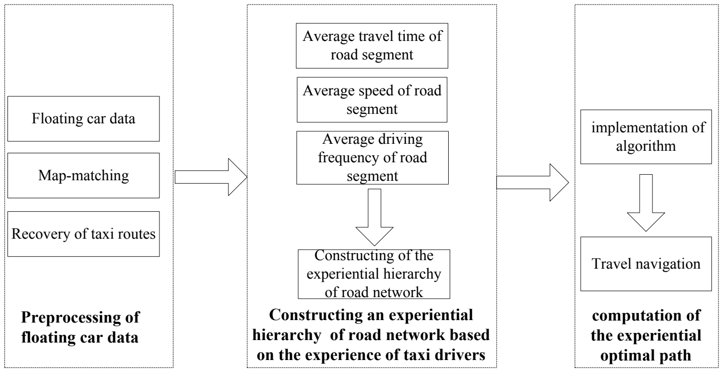

Fig. 1. technical process

3. Extraction of Taxi Routes

3.1. Characteristics of the routes selected by taxi driver

Although the routes selected by each taxi may exist a certain degree of difference, but the routes still have common characteristics. Their characteristics are mainly manifested in the following aspects:

(1) Taxi driver does not select the routes completely in accordance with the shortest route, but avoid some routes with the high probability of congestion.

(2) Taxi driver tends to choose the routes with higher traffic capacity and better ambient environment, and also to minimize the switch between different levels of road.

(3) Taxi driver does not blindly select the high-level roads; however, they select some lower-level roads with better traffic environment. So they can avoid some roads with high probability of congestion and achieve the purpose of reducing travel time.

(4) Taxi driver often prefers to select the road with fewer traffic lights and electronic eye.

These considerations are implicit in a large number of taxi trajectory, so analysis the routes of a large number of taxis and apply it to the path planning can provide a better route for public travel.

3.2. Information table for the taxi traveling track

Since the end of 2005, the taxis in Guangzhou have comprehensively been installed the floating car data acquisition system, which can continually collect data with GPS information.

Data information collected by the floating car generally includes the time, latitude and longitude coordinate, the instantaneous speed of the vehicles and the traveling direction, etc. The original floating car data are tracking points recorded by a GPS receiver installed in a taxi, and they need to be corrected. Map-matching technology must be used to correct the floating car data, then check the validity of the data and filter the data that are errors or missing. However, after the above steps, the data are still based on discrete time information, and therefore can't directly reflect the travel characteristic of taxi, so the data are required for further processing.

When the taxi is no load and carrying with passengers, the characteristics of its behavior are quite different. For example, when the taxi is no load, taxi driver tends to look for passengers in the whole city or relatively fixed urban area, and its travel speed is slower. On the contrary, when the taxi is carrying with passengers, taxi driver has fixed target. It can be predicted that there will be significant differences in the traveling characteristics of the taxi under the two states. Therefore, this article splits the traveling track of taxi into no load and carrying trajectory. Because the real traveling track should be continuous set of road segments, so the GPS data "point" information is needed to be converted to the "line" information.

It can be seen from table 1 that *CarState* is used to identify the state of taxi: carrying with passengers or no load. When *CarState* = 4, it means the taxi is empty, correspondingly, *CarState* = 5 means the taxi is carrying. Fig.2 shows how the state of taxi changes over time. It can be seen from Fig.2 that when the taxi is carrying, its state remains the same for a period of time. So traveling track of taxi can be spliced into no load and carrying trajectory through the transformation of the state, then the taxi's every carrying or no-load trajectory can be extracted.

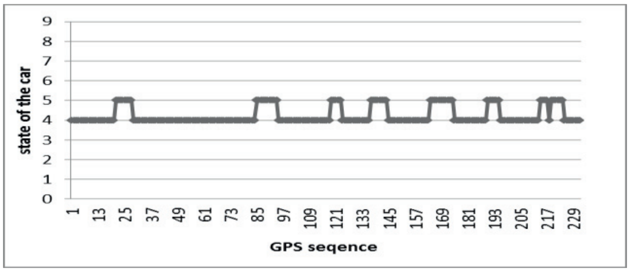

Fig. 2. change of the state

The design of the taxi track information table, as shown in Table 1, includes the starting and ending time, travel time, mileage, average speed, traveling track sequence and so on.

Table 1. Trajectory table of a taxi

Name	Example	Illustration
ID	1925945	Traveling track ID
CarID	1673908319	Vehicle number
FnodeID	3271787741	Start point ID
TnodeID	3272153201	End point ID
FlinkID	2980	Start road segment ID
TlinkID	1739	End road segment ID
CarState	5	State of vehicle(4 or 5)
BeginTime	2011-6-20 6:08:50	Starting time
EndTime	2011-6-20 6:21:08	Ending time
TravelTime	738	Travel time(s)
LinkSeq	213-2700-2702-2703-2712	Sequence of the road segments
Length	6199	Route length（m）
AvgSpeed	30.23	Average speed (km/h)

3.3. Process of the extraction of taxi routes

By the transformation of the state of the vehicle, we can segment the traveling track of the taxi and extract a set of taxi routes that are no load or carrying. The specific process is as follows:

Taxi A starts from the starting point S equipped with passengers at the time t_s, the state flag $CarState$ changes from 4 to 5. When the car reaches its destination at the time t_e, the state flag $CarState$ becomes 4. Thereby, all GPS matched data sequence of taxi A can be obtained from time t_s to time t_e .When the $CarState$ changes from 4 to 5, the routes of the taxi that is no load can be extracted.

Fig.3 is a sequence of taxi trajectories, the road segments in thick line are the sequence of tracks:

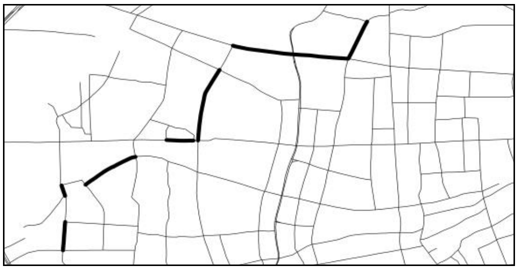

Fig.3 sequence of taxi trajectories

4. Construction of The Experiential Hierarchy of Road Network

4.1. Screening of The Experiential Routes Set

Because the driving behavior of taxi is different when it is empty and carrying with passengers, so the trajectory needs to be screened by the average traveling speed. The route with speed lower than a given threshold value should be excluded. Let V_i be the average speed of route i , V_T be the speed threshold value. The track that is in line with the experiential path filter condition is shown in the following formula:

$$S_{ER} = \{i \mid v_i > V_T\} \tag{1}$$

In different periods, the traffic conditions, the choice of taxi drivers, as well as the travel demand are different. Therefore, it is necessary to set limits at the experience paths in a certain period of time.

Let T be the time interval, so the experience routes within the time interval T are defined by:

$$S'_{ER} = \{i \mid v_i > V_T, t_{is}, t_{im} \in T\} \tag{2}$$

t_{is} is the starting time of the route， t_{im} is the ending time of the route. In any time interval, an experiential route set can be established by the formula (2). However, road traffic is a continuous change process, Fig.4 shows that more than 95% of the travel time is less than half an hour, so setting half an hour as a time interval will make many historical trajectories excluded and affect the subsequent results. Therefore, in order to get a good sample of the data, it is recommended that the time interval is set to a value of not less than one hour.

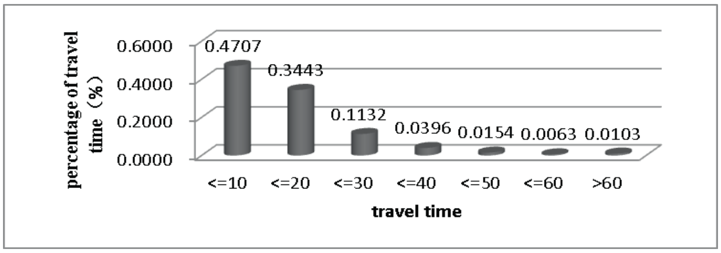

Fig.4 distribution graph of the travel time

4.2. Classification of road network based on the experiential routes set

City roads are generally divided into four grades: express way, main road, secondary road and branch, but this classification of the urban road does not reflect the level of actual traffic conditions. In this paper, an experiential road hierarchy for the network is constructed using experiential routes. An experiential road hierarchy for the network can be constructed based on classification of the travel frequency. The more frequently that road segments are traveled as part of the experiential routes, the more likely the segments are assigned to the upper level road category. An experiential road hierarchy can then be constructed.

Let $n_i(t)$ be the total number of times that the road segments are traveled within the given time interval, then there are:

$$N_i = n_i(t) = \sum_{j=1}^{j=m} n_j^i(t) \tag{3}$$

$$N_{总} = \sum_{i=1}^{i=n} N_i = \sum_{i=1}^{i=n} n_i(t) \tag{4}$$

$n_i(t)$, N_i are the travel times of road segment i within the time interval T, m is the total number of taxis, n is number of road segments.

The larger $n_i(t)$ is, the more likely taxi drivers tend to select road segment i, and vice versa. Then divide the road network into different grades by sorting the value of $n_i(t)$ from high to low. Assuming that all road segments of the urban road network are divided into d grades, g_d is the minimum travel times that meet the level d, H_d is the route set that belongs to the level d, there are:

$$H_d = \{i \mid n_i(t) \geq g_d \ and \ n_i(t) < g_{d+1}\} \tag{5}$$

Assume that the city road network is divided into two grades: the road with high frequency of use (also known as the experiential road layer), and the road with low frequency of use. Jiang(2009) points out that 80% of the urban road network traffic flow is undertaken by 20% of the streets and 80% and 1% of the top streets contribute more than 20% of the traffic flow. This paper uses this study result to determine the division of high frequency road and the low frequency road. As is shown in Fig.5, the abscissa of the graph is the serial number of the road segments ordered by the ratio of the of each road segments to $n_i(t)$ of all road segments. The ordinate is the cumulate percent of the $n_i(t)$ of each road segments and $n_i(t)$ of all road segments. From Fig.5, we can see that it is similar with the results obtained in Jiang(2009) study, that 80% of the taxi travel frequency concentrated in about 20% of the urban road. Therefore, this paper uses it as the method of determining the value of g_2. The specific steps are as follows: Let G be the number of roads in the urban road network and then sort the N_i of all the roads in descending order. So there exists g_2 that the ratio of the sum of all N_i greater than g_2 to the sum of all N_i is approximately 0.8, and g_2 is the critical value that road network is divided into two different levels. N_i of road frequently used is larger than or equal to the value of g_2 and N_i of the road rarely used is smaller than the g_2.

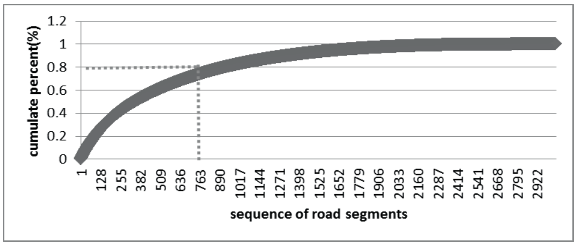

Fig.5 distribution graph of the travel time

5. Hierarchical Path Planning Algorithm Based On Experiential Road Network

Hierarchical path planning algorithm can be briefly described as follows: when given two nodes in the road network as starting point and destination point, first determine whether they fall into the high-level road network, if it is, then directly use the corresponding subroutine in the high-level road network to acquire planning results. Otherwise, you must search in a small area to find the path to the high-level road network. At this time, high-level road network in this area and low-level road network should be both considered; then, just consider the corresponding starting point and destination point in the high-level road network and search the optimal path of them. Finally, combine them and output as a final path planning result. In other words, in this case, final optimal path should include two parts: the path from a start node or destination node in the lower road network to the corresponding starting node or destination node in the high-level road network, the path between the corresponding starting node and destination node in high-level road network.

5.1. Definition and structure of the experiential road network

The experiential road network is constructed using $n_i(t)$. According to the algorithm in this paper, experiential road network is expressed as follows:

$G=(V,E)$ comprises of two sets: a non-empty set of vertices and limited arcs. It meets the conditions of:

(1)The Body of the arcs is Top –k road segments sorted by the $n_i(t)$.

(2)The total travel frequency of the top-k road segments reaches approximately 80% of the sum of travel frequency of all roads.

(3) G remains connectivity.

Hierarchical road network that path planning algorithm bases on requires the following characteristics:

(1) Road network can be divided according to the road grade, traffic volume and running speed.

(2)The details of level gradually increase from high to low, and high-level is a subset of low-level

(3)Each level of the road network is connected.

5.2. Implementation of the algorithm

After constructing an experiential road network, we can apply the hierarchical path planning algorithm to search an optimal path. The algorithm is divided into three steps, as shown in Fig.6 (a): firstly, find the shortest path from S to the entry point S’ and the shortest path from exit point D’ to D. Secondly, find the shortest path between S’ to D’ in the experiential road network layer. Finally, merge the three paths. The algorithm flow is as follows:

Step1.As shown in Fig.6 (a), if the starting point S and D are in the experiential road network layer, then directly search for the shortest path in the experiential road network layer;

Step2.If the starting point S is not in the experiential road network layer, if there are roads of the experiential road network layer in the rectangular region R2, then use the shortest path algorithm to search the entrance point S' of the experiential e road network layer that is nearest to the point S, and record the S-S '. Similarly, search entrance point D' of the ending point D in the range of R2 and record the D-D ', otherwise turn to Step4;

Step3.Search for the shortest path of S' and D' in the experiential road network layer, and combine each partial path to obtain a complete path;

Step4.Use Dijkstra algorithm to search for the shortest path within the R2 in the basic road network.

In the algorithm, passageways of the higher level are found by restricting search area. As shown in Fig.6 (b), firstly, the connection of the S and D is used to form a rectangle R1 as its diagonal, and then R1 is outwardly extended of each side of a threshold value h and form a rectangular R2. R2 is used to determine whether you need to search for the passageways of the experiential road network layer and it is the range of search. The value of h in the algorithm is 0.001 degrees. Algorithm flow chart is shown in Fig.7.

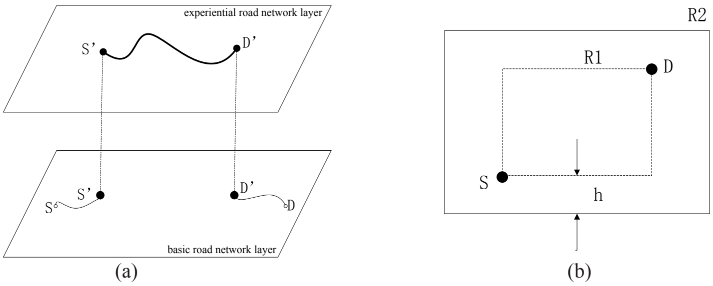

Fig.6 (a)experiential hierarchy of road network; (b)definition of R2

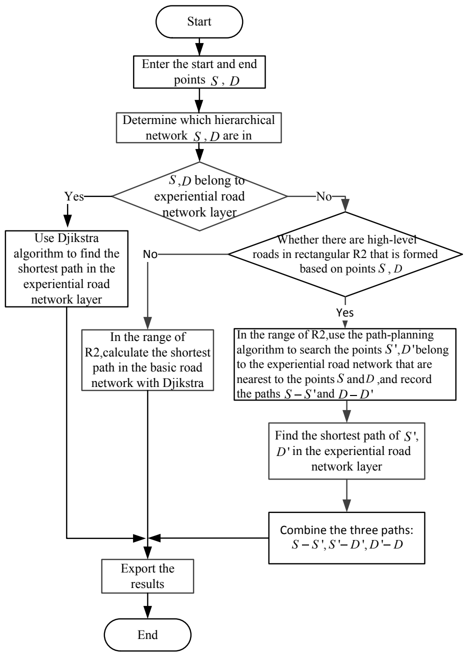

Fig.7 algorithm flow

6. Case Study

This paper selected the electronic map of Guangzhou traffic road as basic experimental data. There are 3042 road segments and 2112 nodes. The taxi GPS data are more than 2900 million effective routes data of over 17,000 taxis in Guangzhou and the GPS data start from June 18 to July 18 in 2011. In this article, due to the small size of the road network, the road network is divided into two classes, respectively, for the road network layer with high frequency of use and the road network layer with low frequency of use. In the experiment, calculating the travel time road needs the speed of the road segments, which is calculated by the average speed with an interval of five minutes that of Guangzhou main road network traffic state analysis and evaluation of the system.

In the experimental analysis chart, SP represents the shortest path planning algorithm, EHP represents path planning algorithm based on the experiential hierarchical road network, RHP represents the path planning algorithm based on the high-level road network formed by the roads with grade 5 and 6 in the urban road network. The three algorithms are all used to compute shortest path between arbitrary OD pairs. so the experiment randomly generated 30 OD pairs and used the above three algorithms to find the different paths.

For the EHP algorithm, because the traffic flow of urban road segments changes over time, the taxi driver’s choice of the road in the flat peak and peak hours will be adjusted accordingly and the road grade division according to the taxi experience will be different. Therefore, it is necessary to consider the time when constructing an experiential network. According to the characteristics and condition of the road traffic of Guangzhou, this article selected four representative time period to establish four experiential road networks, and used the EHP algorithms to find paths in different time periods. The specific time period division is as shown below:

Table 2. period division

Period	Morning peak	Off-peak	Evening peak	Off-peak
T	7:00-9:00	12:00-14:00	17:00-19:00	21:00-24:00

6.1. Comparison of route length

Fig.9 shows a comparison of the route length obtained by the four algorithms, the height of the bar represents the length of the route, and the right side of the vertical axis represents the ratio of the length of the RHP, EHP, SP to the shortest path algorithm. It can be seen from Fig.9 that change range of the ratio of RHP algorithm route length to the SP algorithm route length is big and it is especially obvious when the path is shorter. RHP algorithm tends to select high-grade road, which is prone to detours. It will choose the high-grade road even if the distance between OD is close. The ratio of the route length of the EP algorithm to the route length of the SP algorithm is relatively stable.

Fig.10 is a comparison chart of the three algorithms under 30 ODs on the total length of route in each period, overall, the total route length of the SP algorithm is the shortest, the total route length of RHP algorithm the longest and the EHP algorithm is in between.

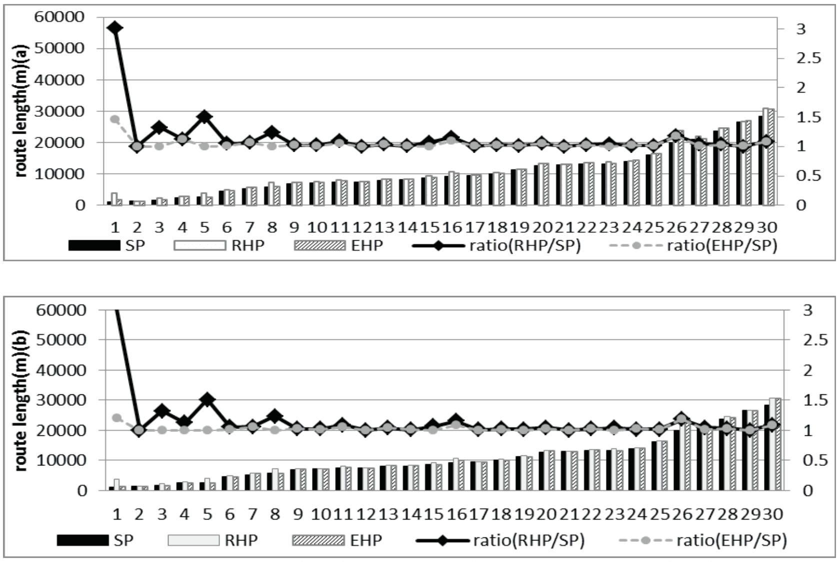

Fig.9. (a) comparison of route length at 7:00–9:00;(b) comparison of route length at 12:00–14:00

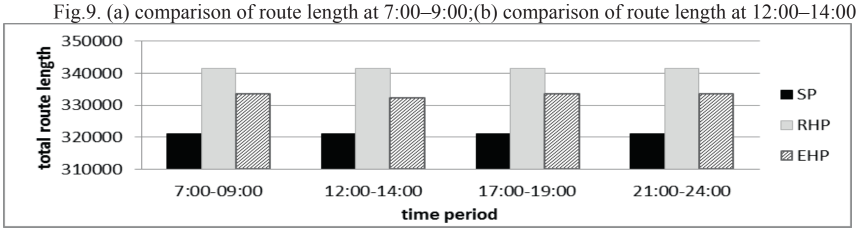

Fig.10 total route length comparison for the three algorithms

6.2. Comparison of travel time

Fig.11 shows the comparison of the travel time of the route calculated by the four algorithms, the height of the bar represents the travel time of the route, and the right side of the vertical axis represents the ratio of the travel time of the RHP, EHP, SP to the shortest path algorithm.

For the EHP algorithm, the results of it are only about 75% less than or equal to the results of the SP algorithm, but travel time of results obtained by the RHP algorithm are only about 66% equal to or less than that of the SP algorithm. Compared with the ratio of the travel time of the RHP algorithm to the shortest path algorithm, the value of the Ratio (RHP / SP) is of lager instability, it has bigger fluctuations when the route length is shorter.

Fig.12 shows that in the overall, travel time of EHP algorithm is the shortest, then it is the RHP algorithm, and the total travel time of SP algorithm is the largest. But the difference between the RHP algorithm and the SP algorithm is not big.

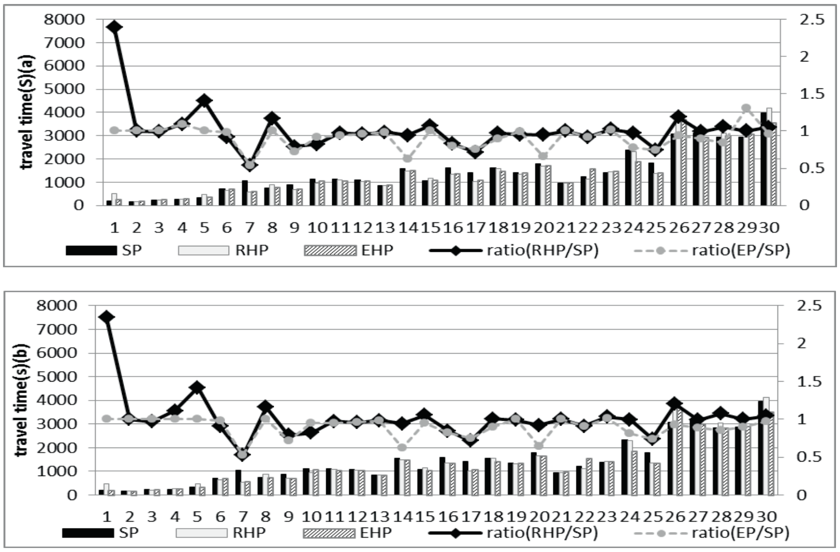

Fig.11.(a)comparison of travel time at 7:00–9:00; (b) comparison of travel time at 12:00–14:00

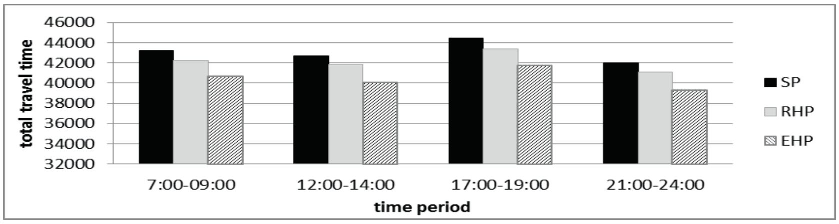

Fig.12 total travel time comparison for three algorithms

7. Conclusions

In this paper, traveling routes of the taxi are extracted combining with the characteristics of the taxi GPS data. Based on the traveling routes of the taxi, each road segment's frequency of travel in different time periods is calculated. Then construct an experiential road network using the experience of taxi driver. Based on the above steps, a hierarchical path planning method using the experience of taxi drivers is proposed. Compared with the classical shortest path planning algorithm, this method has two advantages: first, the use of a hierarchical search strategy based on hierarchical road network can greatly reduce the amount of computation; second, the experiential hierarchy for the road network is a dynamic network, which is able to reflect the real-time traffic and make the planning results more reasonable.

Finally, this paper selects Guangzhou as a study area and randomly generates 30 OD pairs. Then compares the method proposed with the traditional shortest path algorithm and hierarchical algorithm which is based on the natural road network from the two aspects of travel time and route length. The results show that compared with the results of the other two path planning algorithm, the path planned by this method is better with shorter travel time, and can be dynamically adjusted according to the different time periods.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (No. 41271181), the National 863 plans projects（2011AA110306）and the Work safety special fund of Guangdong province in 2011(2011-118).

References

Sacerdoti E D. (1974). Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, 5(2), 115 - 135.

Korf R E.(1987). Planning as search: A quantitative approach. *Artificial Intelligence*, 33(1), 65 - 88.

Jung S, Pramanik S. (2002). An efficient path computation model for hierarchically structured topographical road maps. *Knowledge and Data Engineering, IEEE Transactions on*, 14(5), 1029-1046.

Jagadeesh G, Srikanthan T, Quek K. (2002). Heuristic techniques for accelerating hierarchical routing on road networks. *Intelligent Transportation Systems, IEEE Transactions on*, 3 (4), 301 - 309.

Chou Y L, Romeijn H E, Smith R L. (1998). Approximating shortest paths in large-scale networks with an application to intelligent transportation systems. *INFORMS Journal on Computing*, 10(2), 163 - 179.

Liu B. (1997). Route finding by using knowledge about the road network. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 27(4), 436 - 448.

Car A,Frank A. (1994). General principles of hierarchical spatial reasoning: the case of wayfinding. Scotland: pp. 646 - 664.

WENG M,WU H H,DU Q Y et al(2006). A Heuristic and Hierarchical Wayfinding Algorithm Based on the Knowledge of Road Network. *Geomatics and Information Science of Wuhan University*. 31(4), 360 - 363.

GAO S,LU F. (2009). A Path Finding Heuristic Algorithm for A Road Network Hierarchy. *Geo-Information Science*, 11(2), 151 - 156.

WU X L,LI Q Q,REN F. (2006). Bidirectional A* Algorithm Based On Hierarchical And Block Data Organization. *Journal of Geomatics*, 31(6).

ZHONG H L,ZHANG M,SHI Y Q et al. (2011). Efficient Route Plan Algorithm Based on Multi-level Road Network Strategy. *Journal of Southwest Jiaotong University*, 46(4), 645 - 650.

TANG L L,CHANG X M,LI Q Q. (2010). The Knowledge Modeling and Route Planning Based on Taxi' Experience. *Acta Geodaetica et Cartographica Sinica*, 39(4).

YANG Y. (2009). Optimal Path Planning Problem Under Real-time Traffic Network, Shanghai: Shanghai Jiao Tong University.

Jiang B. (2009). Street hierarchies: a minority of streets account for a majority of traffic flow. *International Journal of Geographical Information Science*, 23(8), 1033 - 1048.