Study on Vehicle Travel Trajectory Completion based on AHP-Entropy-TOPSIS Method

Meimei He ¹, Xianyu Wu ¹*

¹ School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
*Corresponding author’s e-mail: 18120805@bjtu.edu.cn

Abstract. In this paper, we proposed a vehicle travel trajectory completion algorithm based on license plate recognition data. Firstly, a processing method for redundant license plate data is proposed. Secondly, the abnormal travel trajectory is extracted based on the travel time threshold between adjacent nodes. Then, the alternative trajectory was obtained through the K shortest path algorithm. The combined algorithm of AHP and EWM was used to calculate the weight of multi-objective attribute indexes. The Hierarchical-Entropy-TOPSIS method for determining vehicle travel trajectory was constructed to complete vehicle travel trajectory decision-making. Finally, the reliability of the algorithm is verified by the license plate data of Yicheng County, Linfen City.

1. Introduction

A large amount of license plate data contains rich traffic information, and effective use of accurate license plate data can truly restore the traffic characteristics of the actual road network. In practice, due to the limitations of the license plate recognition equipment, the original data often contains incorrect or missing data. Due to the unreasonable location and abnormal operation of the license plate recognition equipment, some intersections in the road network have no data, which makes it difficult to extract the travel trajectory of the vehicle. Therefore, it is of great significance for the effective use of license plate data to complete the data of missing intersection and extract the complete vehicle travel trajectory.

In the early research, license plate recognition data is mostly used in the study of expressway toll collection system. Liang[1] used license plate recognition technology to strengthen the management of urban road and bridge toll stations, effectively intercepting local unpaid annual fee vehicles, and improving the operating efficiency of the toll system. With the development of traffic information collection technology, license plate recognition data is widely used to analyze vehicle travel time. Gong Yue et al.[2] made relevant research on road trip time prediction. In recent years, some scholars have conducted research on how to extract vehicle travel characteristics and traffic state discrimination from license plate recognition data. Chang Yujiao et al.[3] proposed traffic travel characteristics based on license plate recognition data, and used the K-means clustering algorithm to effectively extract commuter vehicles. In applying license plate recognition data to vehicle trajectory research, Yang Shuai et al.[4] established the TOPSISI algorithm with OWA operator, and completed the reconstruction of vehicle travel trajectory based on bayonet data.

Based on previous studies, this paper proposes the application of improved weight calculation TOPSIS method to determine travel trajectory. First, use the K shortest path algorithm to search for the first K trajectories, and then determine the trajectory decision index. Combining the qualitative analysis of the analytic hierarchy process and the quantitative analysis of the entropy weight method, the weight
of the attribute value of the decision index is more objectively determined to make the final trajectory decision. The result is more reasonable. The rest of this paper is organized as follows. Section 2 is related work, including data obtain and pre-processing and extraction of vehicle travel trajectories, Section 3 builds the TOPSIS algorithm based on the improved weight calculation method. Section 4 is case analysis, which verifies the effectiveness of the algorithm.

2. Related work

2.1. Data Obtain and Preprocessing

The study area is Yicheng County, Linfen City, Shanxi Province, which contains 27 intersection detection points. The original data recording time is from 6:00 to 23:59, a total of 1,12067 license plate data. In this paper, we use OpenStreetMap API and ArcGIS to obtain the shapefile of the research area, then the road network topology model of the study area is constructed as shown in Figure 1. Numerically numbered nodes are known intersection detection points in the original data, and letter-numbered nodes are unknown.

![Figure 1. Shapefile of the research area.](image1)

![Figure 2. The road network topology model.](image2)

The classification of raw data and noise data is shown in Figure 3. There are 77381 valid data. The noise data is mainly redundant data, which is reflected in the same license plate being repeatedly identified at the same intersection in a short time.

![Figure 3. Classification of raw data and noise data.](image3)

In order to improve the efficiency of the algorithm, for the above redundant data, this article designs the method to remove duplicate values as follows:

1. Determine the interval time threshold based on the average minimum red light time in Yicheng County, $T=40\text{ s}$;
2. The raw data is arranged in ascending order of time with the license plate number and capture time as the main fields;
3. If the time interval $\Delta t$ for the same car to pass the detection point twice is less than the time threshold $T$, then take the time when the vehicle last passed the inspection point as the available time, and eliminate redundant data;
4. Traverse all license plate numbers until there is no redundant license plate.

2.2. Formatting author names

The trajectory of a vehicle can be defined as a set of two-dimensional sequences composed of points and times in chronological order, expressed as $\text{tr} = \{(p_i, t_i), (p_2, t_2), \ldots, (p_i, t_i)\}$, where $p_i$ is the location of the detection point, $t_i$ is the time when the vehicle passes the inspection point. In this paper, the travel time threshold $B$ is determined based on the ratio of the length of the road segment to the average travel
speed. If the time difference between adjacent travel records is greater than the travel time threshold, the travel trajectory of the vehicle is divided into multiple sub-trajectories.

\[ B = \max \left( \frac{l_j}{v_r} \right) \]  

(1)

Where \( l_j \) is the length of the road section between adjacent nodes, \( v_r \) is the average driving speed.

3. Vehicle Travel trajectory Completion Algorithm

3.1. Problem Description

Due to the lack or abnormality of the license plate recognition equipment, the travel records of the vehicle at some intersections are missing, and the complete travel trajectory cannot be extracted. When the missing intersections between the origin and destination points are greater than or equal to two, which shows as Figure 4, the travel trajectory of the vehicle cannot be complemented by simple trajectory decisions.

3.2. Algorithm Design

To complete the trajectory from O to S, the core is to search for the possible travel trajectory between two points through the shortest path algorithm. However, in actual decision-making, the users not only want to get the best decision, but also want to get the second-best, second-best reference decision. First, we use the K-shortest path algorithm to search the first K paths based on distance. Combined with the actual situation of the road network, K is 5 in this paper. Then the TOPSIS algorithm is used to calculate the relative distance between the alternatives and the ideal solution, and the alternative trajectory solution closest to the ideal solution is selected as the final trajectory. YU and YANG[5], neither of them explained how to calculate the index weight, but just took a fixed value. We propose to use the Analytic Hierarchy Process (AHP) and Entropy Method to comprehensively calculate the index weight, which can achieve both quantitative analysis and qualitative analysis.

3.2.1. Select Decision Index.

This paper considers the characteristics of the road network structure, and selects five indicators of road grade, trajectory length, driving speed, number of vehicle turns, and the number of signalized intersections as indicators for evaluating the importance of road sections. In order to eliminate the influence of different decision-making index dimensions and attributes, standardize decision indicators is necessary, the Min-Max Standardization is as follows:

\[
\xi_{ij} = \frac{\sigma_{ij} - \min \left( \sigma_{ij} \right)}{\max \left( \sigma_{ij} \right) - \min \left( \sigma_{ij} \right)} \quad i \in m, j \in n
\]

(2)

Where \( \xi_{ij} \) represents the standard value of the \( j \)-th decision index attribute on path \( l_j \), \( \max \left( \sigma_{ij} \right) \) and \( \min \left( \sigma_{ij} \right) \) is the maximum and minimum values of the \( j \)-th decision index.

3.2.2. Calculate the Weight Value.
The AHP method quantifies the importance of indicators by constructing a hierarchical structure model based on the scale and judgment principle of pairwise comparison. The decision index weight vector obtained by the analytic hierarchy process is $S_1 = (s_1(1), s_1(2), \cdots, s_1(n))^T$ and satisfied:

$$\sum_{j=1}^{n} s_1(j) = 1$$

$$0 \leq s_1(j) \leq 1$$

The Entropy Method evaluates the degree of variation of index values according to the importance, and uses information entropy to calculate the weight of decision attributes[6]. The information entropy calculation is publicized as follows:

$$e_j = \frac{\sum_{i=1}^{m} z_{ij} \ln z_{ij}}{(\ln m)}$$

$$z_{ij} = \frac{e_j}{\sum_{i=1}^{m} e_j}$$

$$1 \leq j \leq n$$

The decision index weight vector obtained by the Entropy Method is $S_2 = (s_2(1), s_2(2), \cdots, s_2(n))^T$.

$$s_2(j) = \frac{1 - e_j}{n - \sum_{j=1}^{n} e_j}$$ and $$\sum_{j=1}^{n} s_2(j) = 1, 0 \leq s_2(j) \leq 1$$

Based on the principle of minimum information entropy and Lagrange multiplier method, the comprehensive weight vector of the index is calculated $s = (s(1), s(2), \cdots, s(n))^T$.

$$s(j) = \left(\frac{s_1(j) s_2(j)}{1}\right)^\frac{1}{2}$$

$$\sum_{j=1}^{n} \left(\frac{s_1(j) s_2(j)}{1}\right)^\frac{1}{2}$$

$$1 \leq j \leq n$$

3.2.3. Construct a Normalized Decision Matrix.

Combine the comprehensive weight vector $s(j)$ and the standard value of the decision index $\xi_{ij}$ to construct a weighted normalized decision matrix $G$

$$G = (g_{ij})_{m \times n} = (s(j) \cdot \xi_{ij})_{m \times n} \quad i \in m, j \in n$$

3.2.4. Decision Trajectory.

Calculate the difference between the attribute value in the alternative path and the positive and negative ideal solution, we select actual travel time.

$$D_i^+ = \sqrt{\sum_{j=1}^{m} (g_{ij} - g_i^+)^2}$$

$$D_i^- = \sqrt{\sum_{j=1}^{m} (g_{ij} - g_i^-)^2}$$

$$1 \leq i \leq m$$
Where \( g^+_i = \max_{i \in m} g_{ij} \) is the positive ideal solution of the five attribute values in all alternative paths, and \( g^-_i = \min_{i \in m} g_{ij} \) is the negative one.

Calculate \( c_i \) which means how close the candidate trajectory is to the ideal, The largest value is the optimal trajectory.

\[
c_i = \frac{D^-_i}{D^+_i + D^-_i}
\]  

(9)

4. Analysis of Example

Perform the vehicle path extraction, the example vehicle travel trajectory is shown in the table 1.

| Plate number | Time       | Point | Class of plate |
|--------------|------------|-------|----------------|
| Jin A5XXXX  | 2019/11/20 | 1     | 2              |
| ...          |            |       |                |
| Jin A5XXXX  | 2019/11/20 | 11:06:03 | 62        |

Combined with the road network topology, it can be seen that the trajectory of the vehicle from node 1 to node 62 is unknown. There are more than 2 intersections between node 1 and 62. First, according to the K shortest path algorithm, select the first 5 alternative trajectory according to the trajectory length. The trajectory sequence table is shown in the table 2.

| Number | Trajectory Sequence   | Length (km) | Speed (km/h) | Turn | Number of Signalized | Road Grade |
|--------|-----------------------|-------------|--------------|------|----------------------|------------|
| 1      | 1-f-e-c-62            | 2.435       | 22.9         | 1    | 3                    | 0.5        |
| 2      | 1-f-e-d-61-c-62       | 2.640       | 24.8         | 3    | 4                    | 0.46       |
| 3      | 1-f-e-d-61-6-62       | 2.749       | 25.8         | 2    | 5                    | 0.5        |
| 4      | 1-2-a-b-5-6-62        | 2.882       | 27.1         | 2    | 8                    | 0.43       |
| 5      | 1-2-a-d-61-c-62       | 3.188       | 30           | 3    | 5                    | 0.43       |

The weight calculation value is shown in the following table 3.

| Number | AHP  | Entropy Method | Comprehensive weight |
|--------|------|----------------|----------------------|
| 1      | 0.3712 | 0.1568         | 0.2572               |
| 2      | 0.2813 | 0.1689         | 0.2323               |
| 3      | 0.0816 | 0.2273         | 0.1452               |
| 4      | 0.1168 | 0.1514         | 0.1417               |
| 5      | 0.1490 | 0.2954         | 0.2236               |

Substitute the weight value and original index data in the table into the TOPSIS model, use MATLAB software to calculate, get the relative post progress of 5 alternative trajectory, which is \([c_1, c_2, c_3, c_4, c_5]=[0.7675, 0.4651, 0.5536, 0.3268, 0.3938] \). So the first trajectory is the optimal trajectory, but we can see that the value of \( c_i \) has no correlation with the distance. Through the completion of multiple sets of vehicle trajectories, this law can be satisfied, which proves that the TOPSIS algorithm with improved weights is effective.
5. Summary
This paper proposed a TOPSIS trajectory decision algorithm based on analytic hierarchy process and entropy method to calculate the weight and clarified the method of calculating weight, which made the trajectory completion result more authentic. At the same time, the complete vehicle trajectory supplemented by the algorithm in this paper can also be used for road traffic feature analysis, making the analysis results more in line with reality.

References
[1] Liang Lixin. The Application of License Plate Recognition Technology in Toll Stations in Zhongshan City[J]. Guangdong Highway Transportation, 2007(01):56-57.
[2] Gong Yue. Analysis of Traffic Travel Characteristics based on License Plate Recognition Data[D]. Zhejiang University, 2018. p. 31-36.
[3] Chang Yujiao, Yang Dongyuan. Research on Vehicle Identification of Commuting Characteristics based on License Plate Data[J]. Transportation System Engineering and Information, 2016, 16(02): 77-82+112.
[4] Yang Shuai. Research on Vehicle Travel Trajectory Reconstruction Method based on Bayonet Data[C]. China Intelligent Transportation Association. Proceedings of the 11th China Intelligent Transportation Conference. China Intelligent Transportation Association: China Intelligent Transportation Association, 2016:1355-1362.
[5] Haiyang Yu, Vehicle Trajectory Reconstruction from Automatic License Plate Reader Data[J]. International Journal of Distributed Sensor Networks 2018, Vol. 14(2).
[6] Yan Xinxin, Yuan Zhenzhou, Mao Sijie, Wu Yuelin. Slow Traffic and Urban Design Coordination Evaluation Method based on Entropy Weight-TOPSIS Model[J]. Highway and Transportation Science and Technology,2018,35(09):107-114.