Evaluative analysis of traffic guidance forecasting system based on DTW algorithm——With the big data era as the background

Mengyao Tao¹, Chaohong Liu²

¹College of Applied Science and Technology, Hainan University, Danzhou City, 571700, China
²College of Applied Science and Technology, Hainan University, Danzhou City, 571700, China

*Corresponding author’s e-mail: 763@joasu.com

Abstract. With the development of the economy and society and the improvement of transportation infrastructure, the number of motor vehicles and the flow of road vehicles have increased rapidly, resulting in traffic congestion and frequent traffic accidents, which not only affects the efficiency of road traffic, but also affects the realization of social benefits as a whole. Based on this, services such as road condition prediction and real-time road condition detection came into being. In order to study the difference between the predicted value and the real value, this article selects the domestically used AutoNavi map platform and the Baidu map platform as the research objects, and calculates the relevant values. The predicted value provided by the AutoNavi platform is the same as that provided by the Baidu map platform. The real-time detection values were compared, and at the same time, the relevant index in the Baidu map platform was used in the calculation, and the index weight formula was proposed, and the problems in the relevant prediction were found. According to the actual background of the problem, suggestions are made for traffic congestion prediction and smart city construction.

1. Introduction
With the rapid economic development, the number of domestic automobiles has increased very rapidly. As of April 2021, the number of cars in 72 cities across the country has exceeded one million, and the number of motor vehicles in the country is as high as 378 million. The huge number of cars has brought many problems. The rapid increase in demand for road condition forecasting and the development of big data have given birth to road condition forecasting services. Although the service continues to develop and improve, in the face of increasingly complex realities, there is also room for improvement in the traffic diversion forecasting services launched by various platforms.

2. Status Quo of Traffic Forecast
At present, domestic research on big data in traffic congestion is still in its infancy, and there are many other researches in the construction of smart cities, and they are basically based on model predictions. Zhang Lele summarized and discussed the development status and trends of my country’s intelligent transportation system[1]; Yu Zhiqing once proposed the establishment of a traffic visualization platform; Lu Yi, Song Jiaxin and other scholars have conducted applied research on the edge computing of highway intelligent transportation system[2]. There are also related scholars who have
discussed the problem through model establishment. The construction of a smart city requires the improvement of transportation facilities and the reasonable planning of transportation routes. The development of big data has made the past artificial traffic prediction gradually disappear, replaced by traffic monitoring, platform prediction and GPS positioning data collection. However, the existing forecasting methods also have certain problems.

2.1. Flow monitoring has low coverage density and coarser granularity
Traditional traffic monitoring often uses the Microsoft API interface. This method can easily obtain traffic data, but it is restricted by many factors, and the results obtained may be incomplete, granular, and unable to reflect the subtle flow movement process. Traffic monitoring of actual road conditions has a low coverage density. Due to the problems of traffic monitoring, the difficulty of improving the accuracy of prediction is aggravated [3].

2.2. traffic prediction model has too many preset parameters and poor portability
In the past, there were certain technical difficulties in the acquisition, storage and recall of real-time traffic flow, which led to the small volume and type of data for related problems, and the prediction accuracy did not meet certain requirements. Later traditional prediction models are usually converted into mathematical models, with scattered data and low fit.

3. Data Analysis
Although there is still room for improvement in the existing traffic forecasting methods, with the construction of satellites and the development of traffic map platforms, the credibility of the forecast data has also been greatly improved. Nowadays, the Gaode map platform (hereinafter referred to as "AutoNavi") and Baidu map platform (hereinafter referred to as “Baidu”). Choose Jiaochuan toll station, Yinchuan toll station, Dujiaying toll station, Xinghua South toll station, Tuhua toll station, Yongzhen toll station, Yantian toll station, Qingyuan West toll station, Yunlong toll station, Foshan toll station, Pingxiang Toll station, Huanggang toll station, Huangge toll station, Ningbo West toll station (in the following data, the longitude and latitude of the toll station are expressed), the collected data is:

| Latitude and longitude | index | speed1 | status_desc | speed2 | status | congestion_trend |
|------------------------|-------|--------|-------------|--------|--------|------------------|
| 121.646164,29.967735   | 8.38  | 4.9    | 210         | 2.52   | 4      | worse            |
| 106.322049,38.442228   | 6.91  | 6.4    | 520         | 8.06   | 4      | better           |
| 102.79741,24.926086    | 6.44  | 8.5    | 1204        | 17.27  | 4      | worse            |
| 119.85711,32.840296    | 5.25  | 10.7   | 1700        | 8.54   | 4      | same             |
| 113.351729,23.073974   | 5.02  | 9.7    | unknown     | unknown| 1      | better           |
| 118.260327,31.414964   | 4.63  | 17     | unknown     | unknown| 1      | better           |
| 114.225593,22.610626   | 4.26  | 13.6   | 1570        | 14.27  | 3      | worse            |
| 113.029363,23.667142   | 4.18  | 13.8   | 1320        | 19.12  | 3      | better           |
| 121.594327,29.787734   | 4.17  | 10.2   | 1360        | 14.37  | 3      | worse            |
| 121.646164,29.967735   | 4.32  | 9.8    | 210         | 2.01   | 4      | worse            |
| 106.784607,22.092529   | 4.02  | 11     | unknown     | unknown| 1      | better           |
| 116.492412,40.064435   | 3.55  | 23.8   | 1200        | 14.89  | 4      | same             |
| 113.501201,22.853645   | 3.02  | 3.26   | 1940        | 10.96  | 4      | worse            |
| 121.440851,29.816609   | 2.98  | 2.98   | 1500        | 10.07  | 4      | worse            |

a speed1 is 15:48-15:50 Gaode’s estimated vehicle speed km/h
b speed2 is 15:50 Baidu map real-time detection average vehicle speed (km/h)
3.1. Timeliness control and Data collection and weight setting
In order to ensure the timeliness of the data, the research interval of this article is set at 2 minutes. The research time is 15:48. The research data is the 13 most congested toll stations in the "National Toll Station Congestion" at that time. The data search time is less than or equal to the update time of AutoNavi.

The Baidu map platform opens relevant POI data and provides other data except the average speed of the congested road. In order to use various indicators more accurately, after multiple weight design calculations, the formula for determining the indicator weight setting is[4]:

\[(\text{Estimated vehicle speed} \times \text{distance of congestion}) \times 0.5 + \text{section evaluation coefficient} \times 0.3 + \text{prediction of congestion} \times 0.2\]

3.2. Graphic comparison and value setting

Table 2. AutoNavi forecast data and Baidu measured data

![Graph of AutoNavi forecast data and Baidu measured data](image1)

Table 3. After setting the weight data and Baidu measured data

![Graph of weight setting with Baidu real-time detection](image2)

Set pAk to 15:48-15:50, and pBk to 15:50. Baidu map detects the average speed of congested roads in real time, pCk is the estimated vehicle speed calculated by the weight formula.
3.2.1. Calculate the Euclidean distance

Euclidean distance refers to calculating the Euclidean distance of two points corresponding to the trajectory at each time point, and then comprehensively processing the Euclidean distance of all points, including taking the average value, summing, and taking the median.

\[ \text{dist}(p_{Ak}, p_{Bk}) \] represents the distance between users A and B in a certain period of time, \( p_{Ak}, p_{Bk} \) represent the positions of A and B at time k, \( p_{Ak}, x - p_{Bk}, x \) represent the positions of user A and user B in the x dimension, \( p_{Ak}, y - p_{Bk}, y \) represent the positions of users A and B in the y dimension. So the Euclidean distance is:

\[ \text{EU} = \sum_{k=1}^{n} \text{dist}(p_{Ak}, p_{Bk}) = 56.62 \]

In the same way, the Euclidean distance between the \( p_{Bk} \) number set and the \( p_{Ck} \) number set is:

\[ \text{EU} = \sum_{k=1}^{n} \text{dist}(p_{Ck}, p_{Bk}) = 39.0705 \]

3.2.2. DTW fitting

In order to reduce the fitting judgment error due to the limitation of the discrete sampling number of the two trajectory space domains, the DTW algorithm is used to measure the similarity between each point of the two trajectories, and the DTW algorithm measures the fitting degree of \( p_{Ak} \) and \( p_{Bk} \). The evaluation formula is (compared by shortest path):

\[ D(i, j) = \|p_{Ak}, p_{Bkj}\| + \min\{(Di-1, j-1), (Di, j-1), (Di-1, j)\} \]

Among them \( \|p_{Ak}, p_{Bkj}\| \) is the second norm of the coordinates of two points, which is the Euclidean distance between the two points.

\[ \text{DTW}(p_{Ak}, p_{Bk}) = \text{DTW}(k4, k6) = f(m, n) = 178.6354 \]
\[ \text{DTW}(p_{Ck}, p_{Bk}) = f(m, n) = \text{DTW}(k6, k8) = f(m, n) = 41.9005 \]

It can be seen that the formula after the weight setting is the shortest path from k6 to k8. Since the number of objects selected in this experiment is small, the shortest path in the Gaode prediction value is selected again, which is k4 to k6, and the weight setting is calculated at this time. The path after that is 67.526. Under the same position, the formula after the weight setting can get a value with a better fit with the actual value.

4. Results Analysis and Suggestions Outlook

The weight formula set in this article refers to the measurement method and accuracy evaluation of relevant indicators on the Baidu platform. Although the actual measurement value can be more effectively fitted in this experiment, the influence of contingency and other factors cannot be ignored. Interpret the relevant data, and propose special response methods for the accurate prediction of the prediction platform. At the same time, it puts forward relevant suggestions on the improvement of my country's congestion relief mechanism and the construction of smart cities.

4.1. Data and formulas

The weight formula has not yet completely got rid of the poor portability of current forecasting models, but there is still room for improvement in the more advanced platform forecasts. On the one hand, there are large errors in some data, which shows that the two forecasting platforms cannot be completely consistent. The data set with the parameter value of "unknown" in Table 1 is the road section determined by Baidu as "unobstructed", and it can be clearly seen that AutoNavi regards it as a congested road section. Looking back at the relevant latitude and longitude and expanding the distance radius of these positionings, it can be found that the expanded area has a relatively small flow of
people. It can be seen that the precise location prediction should be as small as possible to reduce the scope of the location to reduce the misleading caused by previous information [5]; in the setting of the weight formula, the length of the “congested road section” has a greater influence, and the evaluation of the congested road section and congestion can improve the accuracy of the prediction. The sets of points with a higher degree of fit are generally between k4-k8. Back to the latitude and longitude, and toll. The vicinity of the station is generally near the entrance and exit of the expressway, which is convenient for the estimation of congested road sections. Therefore, to improve the accuracy of the prediction requires attention to the calculation of the length of the congested road section, and the relevant prediction platform can be improved through technology and use the on-board GPS technology to improve the calculation of the congestion distance [6].

4.2. Calculating conditions
The main time background of this investigation is that during the Labor Day holiday in 2021, there are many emergencies and reduced predictability. At the same time, the gathering of vehicles will cause a certain failure of detection technology. During the holidays when the number of people gather is large and concentrated, the past gathering places should be investigated first, and the past population density heat maps should be analyzed to prepare in advance. For the forecast of smaller geographical areas, the positioning system of mobile devices can be used.

4.3. Traffic diversion mechanism and smart city
The construction of smart cities is closely related to the development of big data. Through this special sample study, it is found that there is still much room for improvement and improvement in the current traffic diversion mechanism. Based on the background of the development of big data, traffic forecasting must first focus on the mining, sorting and analysis of big data, and it is necessary to strengthen traffic data detection technology, mobile terminal detection technology, traffic congestion assessment technology, road model refinement technology, geomagnetic technology and other advanced technologies. In addition, the construction of traffic diversion mechanism also needs to be carried out simultaneously with the construction of smart cities, rational use of MaaS technology, building a visualization platform from macro, micro, and meso dimensions, so as to carry out traffic simulation.

5. Conclusion
Based on the background of the booming development of big data, this paper studies the current situation of traffic diversion in the construction of smart cities. The main purpose of using the "weight formula" in this article is not to find a completely accurate forecast value, but hope that through the analysis of related indicators, it is possible to make more detailed and comprehensive suggestions for improving the traffic guidance mechanism. The construction of a smart city can bring many benefits to the lives of residents, but its completion requires not only the increase of welfare facilities, but also the improvement of the urban problem-solving mechanism. The improvement of the traffic diversion mechanism is an important part of the construction of a smart city. Only when the government, technology platform, and residents play their due role together, can the traffic problem be solved at a high level and achieve all-round improvement of social, economic, and ecological benefits.

References
[1] Zhang Lele, Wang Li, Xiao Xiaoling. The development status and trend of intelligent transportation systems in my country. J. Computer Knowledge and Technology, 2021, 17(03): 247-249.
[2] Lu Yi, Song Jiaxin, Zhang Hua. Research on Edge Computing Application of Expressway Intelligent Transportation System.J. Highway, 2021, 66(03): 242-245.
[3] Gong Chen, Shen Yin, Liu Zhaoyong, Yan Xiaoli. Research on the Application of Big Data in Luzhou Intelligent Transportation System in the Information Age. J. Information Technology and Informatization, 2018(09): 180-182.

[4] Lin Lichun, Liu Hua, Hong Dong. Traffic congestion prediction technology based on big data analysis. J. West China Transportation Science and Technology, 2020(09): 138-141.

[5] Xiao Boyan. Research on urban road traffic short-term flow prediction based on big data technology. J. China New Telecommunications, 2020, 22(15): 125-126.

[6] Ding Xiangchao. Application of big data technology in the construction of intelligent transportation platform. J. Computer Knowledge and Technology, 2019, 15(16): 178-179.