



Multi-factor Comprehensive Prediction of Delay Time through Congested Road Sections

Yuhang Wu^{1*} Tong Jiao² Binggang Li³

1. School of Mining Engineering, North China University of Science and Technology, Tangshan, Hebei, 063210, China

2. School of Construction Engineering, North China University of Science and Technology, Tangshan, Hebei, 063210, China

3. School of Information Engineering, North China University of Science and Technology, Tangshan, Hebei, 063210, China

ARTICLE INFO

Article history

Received: 3 September 2020

Revised: 10 September 2020

Accepted: 9 October 2020

Published Online: 16 October 2020

Keywords:

Delay time prediction

Grey correlation analysis

Data fitting

ABSTRACT

The navigation software uses the positioning system to determine the traffic conditions of the road sections in advance, so as to predict the travel time of the road sections. However, in the case of traffic congestion, the accuracy of its prediction time is low. After empirical analysis, this paper establishes a multi-factor synthesis by studying 7 factors: traffic flow, number of stops, traffic light duration, road network density, average speed, road area, and number of intersections the prediction function achieves the purpose of accurately predicting the transit time of congested road sections. The gray correlation coefficients of the seven factors obtained from the gray correlation analysis are: 0.9827, 0.9679, 0.6747, 0.8030, 0.9445, 0.8759, 0.4328. The correlation coefficients of traffic volume, number of stops, average speed, and road congestion delay time were all about 95%, which were the main influencing factors of the study. The prediction needs to be based on functions. This paper fits the main influencing factors to the delay time of congested roads. It is found that the delay time varies parabolically with the traffic flow and the number of stops, and linearly with the average speed. Because the three impact factors have different weights on the delay time of congested roads, demand takes the weight of each factor. Therefore, the gray correlation coefficients occupied by the main influencing factors are normalized to obtain the weights of three of 0.340, 0.334, and 0.326. The weighted fitting function is subjected to nonlinear summation processing to obtain a multi-factor comprehensive prediction function. By comparing the original data with the fitting data and calculating the accuracy of the fitting function, it is found that the accuracy of each fitting function is close to 0, the residual error, the relative error is small, and the accuracy is high.

1. Introduction

In recent years, many countries are developing a real-time monitoring system capable of intelligently detecting traffic conditions. Some of them have achieved

certain results. Regarding the determination of traffic congestion, Andrea et al. Used GPS trackers and smartphones to identify traffic in 2017. Congestion and accidents ^[1], Bauza and Gozalvez detected road traffic congestion based on the information exchange between vehicles and

*Corresponding Author:

Yuhang Wu,

School of Mining Engineering, North China University of Science and Technology, Tangshan, Hebei, 063210, China;

E-mail: 2358198850@qq.com.

vehicles, vehicles and infrastructure nodes in 2013 [2], but after inspection, they found that they have built a section on urban traffic flow. The accuracy of travel time prediction models is not very high. Existing navigation software often uses GPS on taxis installed with this software to obtain real-time data to determine road conditions. However, in actual road conditions, there are many other factors. For example, under the influence of other factors such as road network density, road area, number of intersections, traffic volume, average speed, number of stops, traffic light duration, etc., causing road traffic jams, the speed will be slow, so that the data obtained by people to estimate the speed has a very large error.

In China, many experts and scholars are also engaged in the research of urban road traffic flow and delay forecasting. Among them, Professor He Guoguang of Tianjin University [3] and Professor Yang Zhaosheng of Jilin University [4] have achieved outstanding achievements in this field. They based on multi-dimensional time series and intelligent neural network, a more accurate model for predicting travel time of urban traffic flow sections was established. In addition, Xu Xuemeng, a graduate student of the University of Electronic Science and Technology of China, and her tutor Professor Liu Qiang [5] published intelligent traffic flow prediction and multipath. Research on optimization problems, but the accuracy is not very high when their models are used for large traffic flow densities, mainly due to insufficient consideration of traffic delays in their modeling.

Therefore, this paper takes this factor into consideration, first analyzes in detail the various factors that affect the density of traffic flow, finds seven of the main factors through data processing, and uses it as an index to measure the density of traffic flow. The main purpose is to delay time. Use gray correlation analysis [6] to get the weight of each indicator, and select the three indicators with the highest weighting for analysis, and then use the method of fitting [7] to predict the delay time. Accuracy, combining gray correlation and analytic hierarchy process, to obtain a comprehensive forecast delay time balance function, and substituting other examples into the model to verify the wide applicability and non-uniqueness of the model.

2. Objective Optimization Model Based on Grey Correlation

2.1 Data Processing

According to the 2016 Big Data Report on Intelligent Travel [8], this paper gets the ranking of peak congestion

delay index in the morning and evening on weekdays.

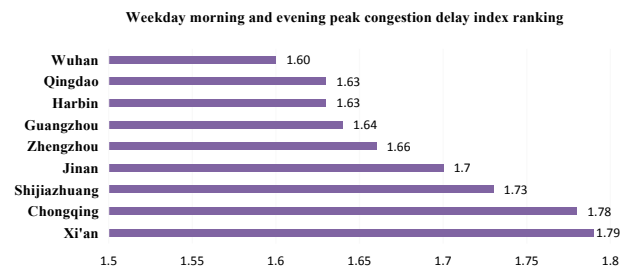


Figure 1. Delayed index ranking during peak periods

Figure 1 shows the morning and evening peak congestion delay time rankings of various cities in the country. It can be seen that the most congested morning and evening peak congestion cities are Xi'an, Chongqing, and Shijiazhuang, and Beijing ranks fourth. In order to make the study representative, this paper chooses the top four cities were selected as the research objects, and the road congestion delay data of each city were randomly selected, and the congested road sections were analyzed in real time.

The common travel modes in life are cars (including taxis), buses, bicycles, and walking. Among them, the impact of each travel mode on the delay time of road trips is shown in Figure 2.

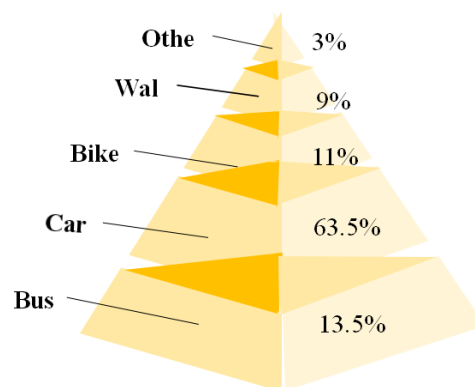


Figure 2. Proportion of travel modes

Figure 2 reflects the proportion of road delays caused by various travel modes. Among them, cars become the main travel mode causing time delays in congested sections, accounting for 63.5% of the total proportion. Therefore, this paper mainly studies the factors that affect the delay time of congested sections and cars. The relationship between the delay times further predicts the delay time of vehicles passing through congested roads.

During the morning and evening peaks of the working day, through the observation of the intersection, the congestion process was found to be five stages: no block, start of block, block, block start to dissipate, and dissipation. . Through statistical analysis, it is found that the morning

and evening peak delays in the four cities are shown in Figure 3.

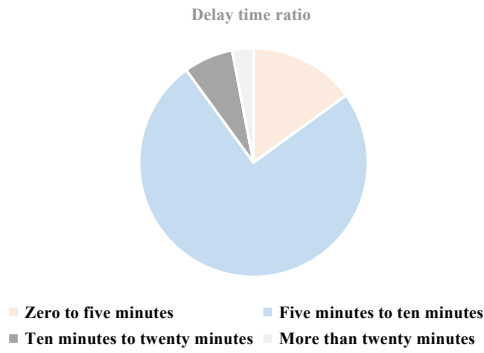


Figure 3. Proportion of delay time

It can be found from Figure 3 that during the morning and evening peak hours of the working day, the driving delay period is concentrated in 5~10 minutes, accounting for 75% of the overall delay time period. Therefore, in this paper, the study is carried out under the condition that the driving delay time is 5~10 minutes.

There are seven factors that affect the delay time during the peak driving period: road network density, road area, number of intersections, traffic volume, average speed, number of stops, and traffic light duration. The research points randomly selected in the four cities are processed and integrated. The data are shown in Table 1 [9].

Table 1. Data integration of factors affecting vehicle delay

Delay time(min)	5.0	7.0	8.0	9.0	10.0
Traffic flow(Vehicle/min)	128.0	170.0	204.0	220.0	248.0
Number of stops(Times/min)	114.0	158.0	172.0	188.0	206.0
Traffic light duration(min)	0.3	0.8	1.0	1.3	1.5
Road network density(article/km ²)	4.0	4.8	5.0	5.6	6.0
Average speed(km/h)	40.0	30.0	25.0	20.0	15.0
Road area(m ²)	120.0	110.0	105.0	100.0	95.0
Number of intersections(article)	6.0	5.0	5.0	4.0	4.0

2.2 Correlation Analysis

Through literature review, this paper finds 7 key factors that influence the delay time through congested road sections: road network density, road area, number of intersections, traffic flow, average speed, number of stops, and traffic light duration. The delay time through the congested section is selected as the reference sequence; the road network density, road area, number of intersections, traffic volume, average speed, number of stops, and traffic light duration are used as comparison sequences, respectively.

$$x_0 = \{x_0^{(k)} | k=1,2,\dots,6\} = (x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(6)})$$

Where k is the point selected by the road segment. Followed by definition ξ_i^{*k} , for comparing the number of correlations between the sequence x_i and the reference sequence x_0 at time K, where $\rho \in (0,1)$ is the resolution factor.

$$\xi_i^{*k} = \frac{\min_s \min_t |x_0(t) - x_s(t)| + \rho \max_s \max_t |x_0(t) - x_s(t)|}{|x_0(k) - x_i(k)| + \rho \max_s \max_t |x_0(t) - x_s(t)|}$$

In the above formula $\min_s \min_t |x_0(t) - x_s(t)|$,

$\rho \max_s \max_t |x_0(t) - x_s(t)|$, Do not refer to the minimum difference and the maximum difference between the series and the associated series. When the resolution coefficient ρ is larger, the resolution is larger, and the smaller the resolution coefficient ρ is, the smaller the resolution

is. Thus, this paper takes $\rho = \frac{\rho}{2} = 0.5$.

Because the correlation coefficient in the formula is an indicator describing the degree of correlation between the comparison sequence and the reference sequence in each year, there are 6 correlation coefficients, so the correlation coefficient is too scattered, which is not conducive to intuitively reflecting the relationship between the comparison sequence and the reference sequence. Therefore, this paper needs to get the average of the correlation coefficient.

$$r_i = \frac{1}{n} \sum_{k=1}^n \xi_i^{*k} \quad k=1,2,3,\dots,7$$

The initialization sequence of each data is substituted into the above formula, and the average correlation coefficients of these 7 sequences can be obtained by using MATLAB software, as shown in Table 3.

Table 2. Relevant factor data

r_1	r_2	r_3	r_4	r_5	r_6	r_7
0.9827	0.9679	0.6747	0.8030	0.9445	0.8759	0.4328

It can be seen from Table 2 that the traffic flow, the number of stops, and the average speed have a great impact on the delay time through congested sections, about 95%. The length of traffic lights, road network density, and the number of intersections have a small effect on delay time through congested sections. Less than 80%, so the delay time and traffic volume, the number of stops, and the average speed through the congested roads are

mainly related to the length of traffic lights, road network density, road area, and the number of intersections.

2.3 Find the Weight of the Main Relevant Factors

Through grey correlation analysis, the traffic volume, the number of stops, and the average speed are selected as the research objects. The correlation coefficients of the three on the delay time through congested roads are 0.9827, 0.9679, and 0.9445.

Therefore, the weights of traffic volume and average number of stops are:

$$w_1 = \frac{0.9827}{0.9827 + 0.9679 + 0.9445} = 0.340$$

$$w_2 = \frac{0.9679}{0.9827 + 0.9679 + 0.9445} = 0.334$$

$$w_3 = \frac{0.9445}{0.9827 + 0.9679 + 0.9445} = 0.326$$

2.4 Perform Data Fitting for Each Factor

By studying the congestion data of Xi'an, Chongqing, Shijiazhuang, and Beijing cars, this paper explores the fitting curves and functions of traffic volume, number of stops, average speed, and road congestion delay time.

After continuous fitting, it is explored that the traffic flow and the delay time through the congested road section change in a quadratic function image. Establish a mathematical fitting model of traffic flow and delay time according to the exploration paper:

$$y = ax^2 + bx + c$$

Parameter estimation observation worthy model:

$$y_i = ax_i^2 + bx_i + c \quad i = 1, 2, \dots, n$$

The accuracy of the fit:

$$Q_i = \sum e_i^2 = \sum (y_i - ax_i^2 - bx_i - cx)^2$$

The parameters can be calculated using MATLAB software:

$$a = -2.222398897e-05 \quad b = 0.049238$$

$$c = -0.86458674168 \quad Q = 0.10530762$$

Get the model function:

$$y_1 = -2.222398897e-05x_1^2 + 0.049238x_1 - 0.86458674168$$

By fitting it, this paper obtains the trend graph of traffic flow and congestion time delay:

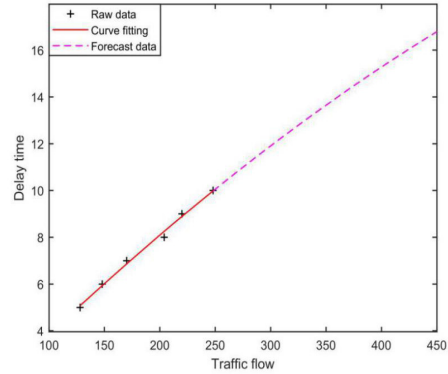


Figure 4. Traffic flow and delay time fit diagram

It can be found from the fitting function that the larger the traffic volume, the greater the probability of congestion and the longer the delay. With more than 250 passing vehicles per minute, road delays can exceed 10 minutes.

By the same reasoning, the fitting curve and function of road congestion delay time, number of stops and average speed are shown in Table 3.

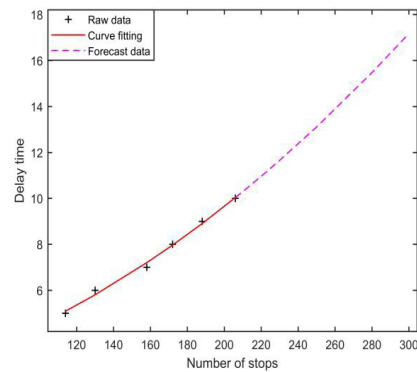


Figure 5. Figure of the number of stops and delays

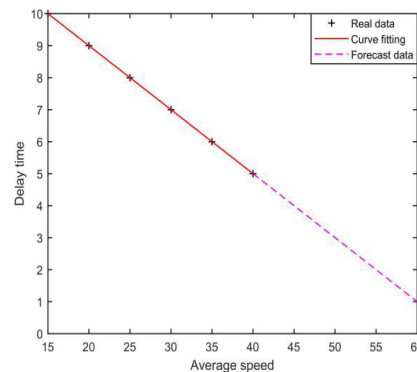


Figure 6. Average speed and delay time fit diagram

Table 3. Delay time and the number of stops and the average speed

Figure of the number of stops and delays	Average speed and delay time fit diagram
Fitting equation $y=1.19264895e-04x_2^2+0.015584332x_2+1.7611$	Fitting equation $y=-0.2x_3+12.996$
Contrast accuracy Q=0.09768167	Contrast accuracy Q=0.00000096

By fitting the curve and function between the two, it can be found that the more times the car stops, the greater the probability of congestion and the longer the delay. When the car stops more than 205 times per minute, the road delay time will exceed 10 minutes; the average speed is inversely proportional to the delay time. The lower the average car speed, the greater the probability of congestion and the longer the delay. When the average speed of a car is below 15 kilometers per hour, road delays can exceed 10 minutes.

2.5 Multi-factor Comprehensive Delay Time Prediction Function

By fitting three main influencing factors to road congestion delay time, this paper obtains three fitting functions about delay time. Because the three factors have different impact weights on road congestion delay time, this paper assigns a weight to the fitting function. A multi-factor prediction function is obtained to comprehensively judge the delay time of congested road sections.

$$y_{\text{Delay}} = 0.340y_1 + 0.334y_2 + 0.326y_3 = 0.340(-2.222398897e-05x_1^2 + 0.049238x_1 - 0.86458674168) + 0.334(1.19264895e-04x_2^2 + 0.015584332x_2 + 1.761123) + 0.326(-0.2x_3 + 12.996)$$

2.6 Model Checking

In order to carry out the simulation test, the real-time data of the research point should be substituted into the comprehensive prediction equation to obtain the fitting value of the delay time. Table 4 shows the residuals, relative errors, and step deviations obtained using real data and fitted data.

Table 4. Lative error, grade ratio deviation value data table

Serial number	Original value	Model value	Residual	Relative error	Ratio deviation
1	5	5.0531	0.0531	0.0106	-0.0425
2	6	5.9110	0.0890	0.0148	-0.0742
3	7	7.0194	0.0194	0.0028	-0.0166
4	8	8.0754	0.0754	0.0094	-0.0656
5	9	8.9307	0.0693	0.0077	-0.0616
6	10	10.0026	0.0026	0.0002	-0.0024

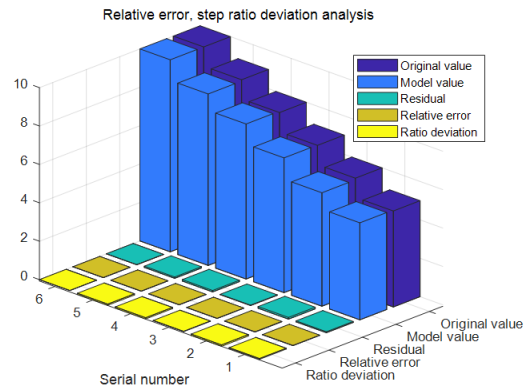


Figure 7. Analysis of relative error and step ratio deviation

After verification, the residual, relative error, and step deviation between the data fitted by the model and the original data are close to 0, which proves that the model has high accuracy and can predict road delays in real time.

3. Conclusion

In this paper, by processing big data of intelligent travel, and combining grey correlation analysis to determine the three factors of traffic flow, number of stops, and average speed play a decisive role in predicting the delay of road congestion. Use the fitting algorithm to process the data of the three main factors to determine the relationship function between each factor and the time required to pass through the congested road. The relationship functions of the three factors are linearly summed to obtain a multi-factor comprehensive prediction function, and the factor data is brought into the multi-factor comprehensive prediction function is used for prediction. It is found that the error of the prediction result is close to 0 through error analysis, which can provide a certain basis for determining the time required to pass through the congested road section.

References

- [1] Eleonora D’Andrea. Francesco Marcelloni. Detection of traffic congestion and incidents from GPS trace_analysis[J/OL]. Expert Systems with Applications, 2017(73):43-56.
- [2] Bauza R, Gozalvez J. Traffic congestion detection in large-scale scenarios using vehicle-to-vehicle communications[J]. Journal of Network and Computer Applications, 2013(05):1295-1307.
- [3] He Guoguang. Study on the Travel Time Budget Method of Vehicle Line Guidance System[J]. Journal of Highway and Transportation, 1998(02):24-27.

- [4] Yang Zhaosheng. Real-time Traffic Flow Artificial Neural Network Prediction Model[J].China Journal of Highway and Transport, 1998(04):89-92.
- [5] Xu Xuemeng. Research on intelligent traffic flow prediction and multipath optimization problem [D]. Chengdu: University of Electronic Science and Technology, 2019.
- [6] Manikyam Sandeep,P. Jamaleswarakumar. Enhancement and gray relational analysis of EN 31 steel using parameters of wire cut electric discharge machine[J]. Materials Today: Proceedings, 2019.
- [7] Su Xiaying.Fitting goodness test of error distribution of linear regression model[D].North China Electric Power University, 2015.
- [8] 2016 Smart Travel Report 0110-v12.2
<https://report.amap.com/predict.do>
- [9] High German Transportation-Transportation Prophet System
<https://report.amap.com/predict.do>
- [10] Zhang Xiaowen.Research on the Guide Signs of Urban Roads in Beijing [J].Highway and Automobile Transportation, 2015(05): 51-54.
- [11] Yang Xingang, Li Desheng, Ren Guoyou. Beijing Road Traffic Accident Forecast Analysis Based on Residual Modified Grey Forecasting Model [J].Mathematics in Practice and Theory, 2018(03): 159-166.
- [12] Tian Yuan. Cancer Metabolism Analysis and Inter-genic Correlation Research Based on Fitted Regression [D]. Jilin: Jilin University, 2018.