Lecture Notes in Electrical Engineering 617

Wuhong Wang
Martin Baumann
Xiaobei Jiang *Editors*

Green, Smart and Connected Transportation Systems

Proceedings of the 9th International Conference on Green Intelligent Transportation Systems and Safety

Springer

Lecture Notes in Electrical Engineering

Volume 617

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Associate Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Executive Editor (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, SCOPUS, MetaPress, Web of Science and Springerlink ****

More information about this series at http://www.springer.com/series/7818

Wuhong Wang · Martin Baumann ·
Xiaobei Jiang
Editors

Green, Smart and Connected Transportation Systems

Proceedings of the 9th International Conference on Green Intelligent Transportation Systems and Safety

Springer

Editors
Wuhong Wang
Beijing Institute of Technology
Beijing, China

Xiaobei Jiang
Beijing Institute of Technology
Beijing, China

Technical University of Munich
Munich, Germany

Martin Baumann
Ulm, Baden-Württemberg, Germany

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-15-0643-7 ISBN 978-981-15-0644-4 (eBook)
https://doi.org/10.1007/978-981-15-0644-4

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Contents

Ordinal Logistic Regression Modeling Research on Decreasing Perceived Metro Transfer Time

Xuesong Feng, Weixin Hua and Xuejun Niu

Abstract This study newly develops two Ordinal Logistic Regression (OLR) models to explore effective ways to save Perceived Transfer Time (PTT) of metro passengers, in view of the difficulty of improving the infrastructure of a metro station. It is found that the PTT will be effectively decreased if the transfer walking congestion is released to be acceptable. Moreover, the congestion on the platform should be eliminated for reducing the PTT. In addition, decreasing the actual transfer waiting time to less than 5.00 min will evidently decrease the PTT. In future works, the effectiveness of the newly developed OLR models needs to be validated in a further and improved by applying them to study the PTT of metro passengers in different cities.

Keywords Perceived transfer time · Perceived transfer waiting time · Ordinal logistic regression · Metro transfer

1 Introduction

Many people today still prefer to utilize private cars in priority, mainly due to unsatisfactory transfers between different public travel routes in their perceptions. It is proved that the convenient transfers at the interchange stations are critical to successfully improving the urban public transport services [1–3] by saving adequate travel time of the passengers [4, 5]. Only in this way can it be possible for more people to abandon private cars, in consideration of their negative attitudes towards making the transfers [2, 6].

Nevertheless, it will be extremely difficult especially for a metro station which is usually constructed underground to objectively reduce the transfer time of its users, due to the restrictions of its infrastructures which are, in fact, impossible to make any

X. Feng (✉) · W. Hua
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
e-mail: xsfeng@bjtu.edu.cn

X. Niu
School of Transportation Management, People's Public Security University of China, Beijing 100038, People's Republic of China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_1

big changes after putting the station into operation. Consequently, in order to improve the transfer services of metro stations, every effort ought to be made to reduce the negative influences upon the Perceived Transfer Time (PTT) of the passengers or even make the passengers underestimate their PTT. To this end, in view of the difference between the time consumed in reality and the perceived time cost [7], distinct effect of various factors on the perceived time costs for metro transfers has to be analyzed in advance.

Though, the perceived time for travels has been studied for a long time, specific analyses on the perceptions of various time costs expensed in public transports are fewer in contrast [8]. For example, González et al. [8] focus on the perceived travel time in tram and find that it is a function of both the commuter characteristics and the time consumed in other travel stages. Fan et al. [9] conduct a passenger survey and record the waiting passengers at different transit stops in videos to compare the perceived and actual waiting time, and discover that basic amenities at transit stops obviously reduce the perceived waiting time. In the limited research on the perceived time of the public transport passengers, studies on the perceived time consumed in metro transfers are especially inadequate. As a result, based on the sample survey on the PTT of the metro passengers in Beijing, two Ordinal Logistic Regression (OLR) models are developed in this work to analyze the effect of different actions on reducing the PTT for the improvement of the metro transfer services.

The latter parts are organized as follows. The survey on the time costs for the metro transfers in Beijing is introduced in Sect. 2, and there is also a preliminary analysis of the survey data. Thereafter, Sect. 3 develops two OLR models to interpret the cumulative probabilities of the PTT and the Perceived Transfer Waiting Time (PTWT), respectively. These two newly developed OLR models are applied in Sect. 4 to predict the probabilistic changes of the PTT for different scenarios taking various actions for the improvement of the transfer service. At last, Sect. 5 makes conclusions, suggests ways to reduce the PTT of metro passengers, and discusses some future research issues.

2 Data Survey and Analysis

The survey on the PTT of the metro passengers was made from November 29th, 2017 to December 5th, 2017, at 4 locations (i.e., Xizhimen Capita Mall, Zhonguancun Plaza Shopping Mall, Xidan Joy City and Xihongmen LIVAT Shopping Center) which are next to each of 4 metro stations in Beijing, respectively. Questionnaires were issued to people staying in the rest areas of these 4 locations after getting their permissions and ensuring that they had the metro transfer experience within one week. There are 3 kinds of the questions on a questionnaire. The first kind of the questions involve the age, educational background, occupation, income, frequency of travels by metro, etc. of a respondent. The second kind of the questions ask the people to tell the details of their last travels by metro, such as travel purpose, travel time period, familiarity with the transfer route, total time cost for the entire trip, and

Table 1 Paired T-tests in perceived and actual groups of samples

Difference	Mean	Standard deviation	T	Sig.	95% confidence interval	
					Lower limit	Upper limit
Perceived walking time − Actual walking time	0.63	0.92	14.85	0.00	0.55	0.72
Perceived walking distance − Actual walking distance	0.86	1.47	12.63	0.00	0.73	0.99
Perceived waiting time − Actual waiting time	0.35	0.69	10.93	0.00	0.29	0.41

so on. The third kind of the questions focus on the perceptions of people about the transfers in their last metro travels, including perceived transfer time consumption, perceived walking distance, perceived waiting time, etc. A total of 490 questionnaires have been distributed and collected. 467 valid questionnaires are obtained finally. In order to compare the differences between the perceptions of the surveyed people and the facts, the field investigations on walking distances, waiting time, walking time, etc. for transfers in reality were carried out immediately after the survey. The paired T-tests on the differences of the perceived and actual walking distances, waiting time and walking time are conducted and the results are presented in Table 1.

According to the mean values of the differences shown in Table 1, it is confirmed that the walking distances, waiting time and walking time in the metro transfers are all generally overestimated by most of the passengers. Moreover, in comparison to the over-perceived time expensed in the waiting stage of a transfer, the overestimate of the time spent in the walking stage is more serious because the time information is ordinarily provided on the platform rather than along the transfer routes. Furthermore, in contrast to relatively easy ways to get the time information, seldom distance information can be obtained usually. Additionally affected together by many other factors including congestion, number of utilized stairs, etc. in transfers, the passengers generally overestimate their walking distances the most in comparison to overestimating the time costs.

3 Modeling Study

The OLR analysis [10, 11] is utilized in the modeling work of this research. As explained by Eq. (1), an OLR model is first established to explain the relationships between the cumulative classification probabilities of the PTT and its correlated variables which are discretized in this work. Moreover, as one of the most important factors of the PTT, the PTWT, which is classified in this research, is also probabilistically explained by its related discretized variables in Eq. (2). These two OLR models

Table 2 Calibration of Eq. (1)

Parameter	Estimated value	Standard deviation	Wald	Sig.	95% confidence interval	
					Lower limit	Upper limit
α_1^{TrT}	−0.40	1.20	0.11	0.74	−2.75	1.96
α_2^{TrT}	2.57	1.21	4.54	0.03	0.21	4.94
α_3^{TrT}	5.14	1.24	17.18	0.00	2.71	7.57
β^{Ge^r}	−0.54	0.19	7.64	0.01	−0.92	−0.16
$\beta_1^{T-En^p}$	−1.16	0.50	5.27	0.02	−2.14	−0.17
$\beta_2^{T-En^p}$	−0.77	0.50	2.36	0.12	−1.76	0.21
β^{T-Wk^p}	0.77	0.33	5.31	0.02	0.11	1.42
$\beta_1^{D-Wk^r}$	−1.10	0.46	5.86	0.02	−1.99	−0.21
$\beta_2^{D-Wk^r}$	−0.95	0.44	4.75	0.03	−1.80	−0.10
$\beta_3^{D-Wk^r}$	−0.25	0.43	0.34	0.56	−1.08	0.59
$\beta_1^{T-Wt^p}$	2.59	1.04	6.24	0.01	0.56	4.62
$\beta_2^{T-Wt^p}$	2.93	0.79	13.75	0.00	1.38	4.48
$\beta_1^{St^r}$	0.82	0.30	7.47	0.01	0.23	1.41
$\beta_2^{St^r}$	−3.48	0.91	14.73	0.00	−5.26	−1.71
β^{Co-Wk^p}	−2.98	0.84	12.45	0.00	−4.63	−1.32
β_1^{WW}	1.23	0.40	9.30	0.00	0.44	2.02
β_2^{WW}	1.00	0.36	7.73	0.01	0.30	1.71
β_1^{SC}	−0.83	0.44	3.63	0.06	−1.69	0.02
β_2^{SC}	−1.00	0.33	9.20	0.00	−1.65	−0.35

work in cooperation to explain the changes of the PTT and explore the effective ways to reduce the PTT of the passengers for the improvement of the metro transfer services. As shown in Tables 2 and 3, the calibration results of both of these two OLR models are acceptable from the perspective of statistics. Moreover, with the respective significances of 0.95 and 0.21, they have also passed their Parallel Line tests.

$$\ln\left(\frac{\gamma_i^{TrT}}{1-\gamma_i^{TrT}}\right) = \alpha_i^{TrT} - \left(\begin{array}{l} \beta^{Ge^r} \times Ge^r + \sum_{j=1}^{n^{EnT}-1} \beta_j^{T-En^p} \times T_j^{En^p} + \beta^{T-Wk^p} \times T^{Wk^p} \\ + \sum_{j=1}^{n^{WkD}-1} \beta_j^{D-Wk^r} \times D_j^{Wk^r} + \sum_{j=1}^{n^{WtT}-1} \beta_j^{T-Wt^p} \times T_j^{Wt^p} + \sum_{j=1}^{n^{St}-1} \beta_j^{St^r} \times St_j^r \\ + \beta^{Co-Wk^p} \times Co^{Wk^p} + \sum_{j=1}^{n^{WtT}-1} \beta_j^{WW} \times T_j^{Wt^p} \times T^{Wk^p} + \sum_{j=1}^{n^{St}-1} \beta_j^{SC} \times St_j^r \times Co^{Wk^p} \end{array}\right) + \varepsilon^{TrT} \quad (1)$$

where,

γ_i^{TrT} Cumulative probability of all the PTT costs belonging to the classes no bigger than Class i, Unit: %,

α_i^{TrT} Constant term of the OLR model for all the PTT costs belonging to the classes no bigger than Class i,

Ge^r 0–1 variable denoting the gender of a passenger making a transfer is male (i.e., 1) or female (i.e., 0),

$T_j^{En^p}$ 0–1 variable representing if the perceived time consumption for an entire trip is at Level j (i.e., 1) or not (i.e., 0),

n^{EnT} Number of all the levels for the perceived time costs of entire trips,

T^{Wk^p} Categorical variable explaining various perceived transfer walking time,

$D_j^{Wk^r}$ 0–1 variable which denotes if the actual transfer walking distance is at Level j (i.e., 1) or not (i.e., 0),

n^{WkD} Number of all the levels for the transfer walking distances in reality,

$T_j^{Wt^p}$ 0–1 variable which indicates whether the PTWT is at Level j (i.e., 1) or not (i.e., 0),

n^{WtT} Number of all the levels for the PTWT,

St_j^r 0–1 variable denoting if the quantity of all the stairs utilized in a transfer is at Level j (i.e., 1) or not (i.e., 0),

n^{St} Number of all the levels for the quantities of the utilized stairs,

Co^{Wk^p} Categorical variable interpreting different perceived congestion degrees for waking in a transfer, and

ε^{TrT} Error term for the cumulative probability of the PTT. It is assumed in this research that ε^{TrT} follows the distribution of $N(0, \sigma_1^2)$.

Table 3 Calibration of Eq. (2)

Parameter	Estimated value	Standard deviation	Wald	Sig.	95% confidence interval	
					Lower limit	Upper limit
α_1^{WtT}	2.68	0.77	12.21	0.00	1.18	4.18
α_2^{WtT}	5.93	0.81	53.28	0.00	4.33	7.52
θ^{T-Wk^p}	1.06	0.31	11.63	0.00	0.45	1.67
$\theta_1^{D-Wk^r}$	−0.81	0.38	4.49	0.03	−1.56	−0.06
$\theta_2^{D-Wk^r}$	−0.98	0.37	6.97	0.01	−1.70	−0.25
$\theta_3^{D-Wk^r}$	−0.65	0.38	2.88	0.09	−1.41	0.10
θ^{T-Wt^r}	2.70	0.44	37.18	0.00	1.84	3.57
$\theta_1^{Co-Pt^p}$	−1.13	0.38	8.78	0.00	−1.87	−0.38
$\theta_2^{Co-Pt^p}$	−0.21	0.21	1.06	0.30	−0.62	0.19
θ^{WW}	−0.56	0.20	7.89	0.00	−0.96	−0.17

$$\ln\left(\frac{\gamma_k^{WtT}}{1-\gamma_k^{WtT}}\right) = \alpha_k^{WtT} - \left(\begin{array}{l}\theta^{T-Wk^p} \times T^{Wk^p} + \sum_{j=1}^{n^{WkD}-1} \theta_j^{D-Wk^r} \times D_j^{Wk^r} + \theta^{T-Wt^r} \times T^{Wt^r} \\ + \sum_{j=1}^{n^{PtCo}-1} \theta_j^{Co-Pt^p} \times Co_j^{Pt^p} + \theta^{WW} \times T^{Wk^p} \times T^{Wt^r}\end{array}\right) + \varepsilon^{WtT} \quad (2)$$

where,

γ_k^{WtT} Cumulative probability of all the PTWT costs belonging to the classes no bigger than Class k, Unit: %,

α_k^{WtT} Constant term of the OLR model for all the PTWT costs belonging to the classes no bigger than Class k,

T^{Wt^r} Categorical variable interpreting different actual transfer waiting time,

$Co_j^{Pt^p}$ 0–1 variable indicating whether the perceived congestion degree on the platform for boarding a train is at Level j (i.e., 1) or not (i.e., 0), and

ε^{WtT} Error term for the cumulative probability of the PTWT. In this study, the distribution of ε^{WtT} is hypothesized to be $N(0, \sigma_2^2)$.

If the predicted classification probability of a PTT cost for a certain class is the biggest in comparison to its predicted probabilities for other classes, the PTT cost is predicted to belong to this class. The surveyed and predicted classes of an accurately predicted sample are the same with each other. Based on this principle, the accuracy of the PTT OLR model explained by Eq. (1) is evaluated to be about 62.10%. In contrast, in spite of its a little lower accuracy (i.e., approximately 59.53%), the PTWT OLR model interpreted by Eq. (2) is still able to predict the classes of the PTWT.

4 Scenario Analysis

Based on both of the calibrated OLR models developed in the previous section, 2 scenarios are studied in this section to analyze their effect on reducing the PTT for the improvement of the metro transfer services in Beijing, in consideration of the practicalities of different actions. Various actions taken in each of the scenarios are explained as follows.

Scenario 1 Release walking congestion. That is, decrease the categorical value of the perceived walking congestion degree of the passengers.

Scenario 2 Release waiting congestion. That is, decrease the perceived waiting congestion degree of the passengers.

It is found that, if the passengers who perceive that the transfer walking environment is very congested are successfully convinced that the congestion decreases to an acceptable degree or disappears, many of them will believe that their transfer time cost is reduced. It is indicated that the release of serious transfer walking congestion

to an acceptable degree will effectively reduce the PTT of some passengers, in consideration of lots of difficulties and relatively poor effect of eliminating the transfer walking congestion. Moreover, if the passengers who are confident that the platform for boarding trains is very congested perceive that the congestion becomes acceptable, their PTT has no change. Only if the congestion on the platform disappears, an apparent share of such passengers will believe that their PTT is reduced. Therefore, it is confirmed that the complete elimination of the congestion on the platform is very important for not only the safety of the passengers on the platform but also the reduction of their PTT.

5 Conclusions

Two OLR models are developed to explore the effective ways to reduce the PTT of the passengers. It is confirmed that the newly developed models are able to rationally interpret the probabilistic changes of the PTT with its determining factors. It is found that the PTT of the passengers will be effectively decreased as long as the serious congestion during walking in transfer becomes acceptable rather than eliminated. Eliminating the congestion on the platform is essential to the valid decrease of the PTT. Moreover, reducing the transfer waiting time of the passengers to less than 5.00 min in reality will evidently decrease the PTT. In future research, the newly developed OLR models needs to be used to analyze the PTT of the metro passengers in different cities to further validate and improve their effectiveness.

Acknowledgements This study is supported by National Natural Science Foundation of China [grant number 71571011] and the Fundamental Research Funds for the Central Universities [grant number 2018JBM022].

References

1. Hernandez S, Monzon A (2016) Key factors for defining an efficient urban transport interchange: users' perceptions. Cities 50:158–167
2. Silva JA, Bazrafshan H (2013) User satisfaction of intermodal transfer facilities in Lisbon, Portugal: analysis with structural equations modeling. Transp Res Rec J Transp Res Board 2350:102–110
3. Iseki H, Taylor BD (2010) Style versus service? An analysis of user perceptions of transit stops and stations. J Public Transp 13:23–48
4. Chowdhury S, Ceder A, Schwalger B (2015) The effects of travel time and cost savings on commuters' decision to travel on public transport routes involving transfers. J Transp Geogr 43:151–159
5. Bak M, Borkowski P, Pawlowska B (2012) Types of solutions improving passenger transport interconnectivity. Transp Probl 7(1):27–36
6. Guo Z, Wilson NHM (2004) Assessment of the transfer penalty for transit trips: geographic information system-based disaggregate modeling approach. Transp Res Rec J Transp Res Board 1872:10–18

7. Varotto SF, Glerum A, Stathopoulos A et al (2017) Mitigating the impact of errors in travel time reporting on mode choice modelling. J Transp Geogr 62:236–246
8. González RM, Martínez-Budría E, Díaz-Hernández JJ et al (2015) Explanatory factors of distorted perceptions of travel time in tram. Transp Res Part F Traffic Psychol Behav 30:107–114
9. Fan Y, Guthrie A, Levinson D (2016) Waiting time perceptions at transit stops and stations: Effects of basic amenities, gender, and security. Transp Res Part A Policy Practice 88:251–264
10. Kremelberg D (2010) Practical statistics: a quick and easy guide to SPSS, STATA, and other statistical software. Sage Publications Inc., Thousand Oaks
11. Borooah VK (2002) Logit and probit: ordered and multinomial models. Sage Publications Inc., Thousand Oaks

Research on the Site Selection of Distribution Center Based on Centroid Method and Fuzzy Evaluation Method

Mingtao Chen, Chun Bao, Hua Yang and Zhiyuan Wang

Abstract In order to help logistics enterprise to improve the distribution condition, reduce the waste caused when a vehicle is fragmented transporting commodity, optimize logistics system, reduce costs, improve service levels, is conducive to the strategic development of the enterprise. This paper uses the Centroid Method to get some alternative logistics distribution center address, at the same time, considering the various factors influencing the site selection, mainly includes land condition, natural conditions, traffic conditions, operation conditions and the influence of the policies and regulations, uses Fuzzy Evaluation Method to decide the best distribution center. Finally, the effectiveness of the algorithm was verified by the 27 Eurasian supermarket chains in Changchun city, and the address of the best logistics distribution center was obtained.

Keywords Logistics · Centroid method · Fuzzy evaluation method · Best logistics distribution center

1 Introduction

Distribution center is the center of concentration and divergence of goods which can arrange the transportation of goods scientifically and orderly and achieve scale effect and improve service level. It can also avoid overlapping routes and reduce costs. Site selection refers to the management activities that determine the construction, expansion, and consolidation of a physical entity in order to exploit the market and increase productivity or service levels. In general, location method can be divided into two categories with the goal of minimizing transportation costs or maximizing profits.

Research abroad began in 1909 when Alfred Weber [1] fist propose the problem of Weber. He believed that it can deal with the location problem of single warehouse

M. Chen (✉) · C. Bao · H. Yang · Z. Wang
Hualan Design & Consulting Group, NO. 39, HuaDong Road,
Nanning, Guangxi 530000, China
e-mail: cmt@gxhl.com.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_2

to realize the minimum between warehouse and customer. Then Johann Von Thunen [2] and others added the opinions. Hakims [3] propose the problem of p-median and P-center. Since then, the research on site selection had reached a new level. Methods such as Centroid method, CELP method, Baumol-Volvo method [4], P-intermediate problem model [5], Kuehn-Hamburger [6] model. In the 80s Sophrer et al. [7] used the method of weight analysis to discuss the site selection of a factory with qualitative and quantitative analysis.

The exploration of site selection in China began in the 1990s. Gao [8] optimized the Baumol-Volvo method to solve the site selection of logistics Center with Heuristic Algorithm. Li [9] clarified the nonlinear relationship between the demand and the optimal inventory cost of an individual distribution center and then solved it with Genetic Algorithms. Ding et al. [10] conducted 0–1 integer programming which used the standard of the lowest cost, thus make the best choose. Yang [11] establishes a new site selection model for site selection of individual distribution centers for individual species when the demand is a non-negative random variable. Wang et al. [12] used Centroid to conduct site selection and used the moving average method to get the best location of the distribution center to avoid the randomness of site selection. Wang et al. [13] improved the deficiencies of Centroid method and modeling and calculation with dimensional analysis. Zhao et al. [14] used Centroid method to carry out the initial site selection, and then solved through the discrete model to obtain the specific location. Zhao [15] combined with the Centroid method and fuzzy clustering method to reduce the computational intensity.

Therefore, this study combines the Centroid method with the fuzzy evaluation method. Firstly, obtain several alternative distribution center points with Centroid method. At the same time, according to the evaluator's subjective rating to choose the best site of distribution center with qualitative and quantitative fuzzy evaluation method. It has a strong practical value.

2 Definition of Distribution Center

Distribution is defined as "In the economic reasonable area, according to the user's requirements, the goods are chosen, processed, packaged, segmented and matched, and delivered to the designated location on time." which in "Logistics terms" (GB/T 18354-2006).

With the development of logistics, the warehouse with circulation function has gradually evolved into the distribution center. Due to the different development and time, the specific definition of distribution center is not uniform. Japan emphasizes from the supplier's point of view, the distribution center is to meet the customer requirements of the circulation warehouse. The distribution center is mainly responsible for the specific customers in the small range of the end, with diversified distribution functions as the main, followed by the warehousing function, the information network coverage is comprehensive, and the goods are rich in variety. In short, the distribution center integrates procurement, storage, distribution processing, tally,

delivery and information processing, which is a collection center, distribution center and processing center.

3 Design of Location Model Algorithm

Centroid Method is the most widely used because of its simple operation, low cost and high degree of freedom. Centroid Method is used to determine the location of the facility by solving the center of gravity of discrete particle which suitable for addressing the location of continuous single facility. The goal is to meet the delivery requirements of various demand points and the overall cost is the lowest when implementing Centroid Method. Therefore, it is necessary to obtain accurate information of each demand point before calculation.

3.1 Centroid Method

1. To simplify the calculation, the following assumptions need to be made:

(1) In a certain area, the demand is concentrated at a certain point, and this data is substituted into the calculation, which can represent the demand of a large number of scattered.
(2) Transportation cost is the only evaluation indicator, regardless of the difference between construction cost and operating cost in different regions.
(3) Transportation costs are only proportional to distance, ignoring variable costs and traffic conditions.
(4) The distribution center and each demand point are connected by a straight line, ignoring the actual route.
(5) The decision-making environment is static. In fact, the income and cost of enterprise operation are constantly changing with the change of environment.
(6) The gap between the demand point and the distribution center is much larger than the gap between the supply point and the distribution center, so the latter is ignored.
(7) Demand and location of each demand point are determined and known.

2. Establish Centroid Method

Fist, Suppose the distribution center is built at $L_k(M_k, N_k)$, the various demand points are $L_1, L_2, L_3 \ldots, L_i$ $(i = 1, 2, \ldots, n)$. Assuming the coordinates of each demand point is (M_i, N_i), the quantity demanded is x_i, unit transport rate is f_i. The distance from L_k to each demand point is j_i:

$$j_i = \sqrt{(M_k - M_i)^2 + (N_k - N_i)^2} \tag{1}$$

The freight from the distribution center to various demand points is c_i and total transportation costs is C:

$$c_i = x_i \times j_i \times f_i \tag{2}$$

$$C = \sum_{i=1}^{n} x_i \times j_i \times f_i = \sum_{i=1}^{n} c_i \tag{3}$$

Substituting (1) and (2) into (3):

$$C = \sum_{i=1}^{n} x_i \times f_i \times \sqrt{(M_k - M_i)^2 + (N_k - N_i)^2} \tag{4}$$

The question is that when (M_k, N_k) is equal to what, C is the smallest. At this point, B is the best location for distribution center. Reach the derivative of M_i, N_i through Formula (4). And make the derivative equal to 0.

$$\frac{\partial C}{\partial M_i} = \sum_{i=1}^{n} x_i \times f_i \times (M_k - M_i)/j_i = 0 \tag{5}$$

$$\frac{\partial C}{\partial N_i} = \sum_{i=1}^{n} x_i \times f_i \times (N_k - N_i)/j_i = 0 \tag{6}$$

Calculate the distribution center coordinate:

$$M_k = \frac{\sum_{i=1}^{n} x_i \times f_i \times M_i/j_i}{\sum_{i=1}^{n} x_i \times f_i/j_i} \tag{7}$$

$$N_k = \frac{\sum_{i=1}^{n} x_i \times f_i \times N_i/j_i}{\sum_{i=1}^{n} x_i \times f_i/j_i} \tag{8}$$

ji is unknown, It can be seen from Formula (1), it cannot realize the distance between the distribution center and each demand point if L_k is undetermined. Therefore, the iterative method is used for calculation.

(1) Select the initial location $L_0(M_0, N_0)$ by solving the coordinates of the point of gravity of each demand point.
(2) Calculation the distance *ji* and the total distribution cost C_0. Then calculate the new position $L_1(M_1, N_1)$ and the new total distribution cost C_0.
(3) Comparing C_0. with C_1, if $C_0 < C_1$ then L_0 is the distribution center location. But if $C_1 < C_0$, repeat the calculation of Formula (3) until $C_{k+1} < C_k$.

3.2 Fuzzy Evaluation Method

Centroid Method is to match the goal of the lowest total distribution cost. This method does not fully consider other factors such as natural environment, human cost, land conditions, etc. Therefore we can use Centroid Method to find a optimal point in the large range on the macro level. Then determine the optimal location of distribution center by fuzzy evaluation method on the micro level.

In the application, use fuzzy evaluation method to take into account the relationship between the various factors which considered both the fixed cost and the variable cost. The process is as follows:

(1) Select the alternative plot with numerous influencing factors; (2) Classify and stratify the influence factors of site selection. (3) Assign weights and values. (4) Calculate and compare by the mathematical method of fuzzy evaluation and finally obtain the most suitable location selection scheme.

3.2.1 Factors to Consider When Applying the Fuzzy Evaluation Method

(1) Land Conditions: The price, area, terrain, and infrastructure of the alternative land.

Land prices are the key influencing factors for the location selection of logistics facilities such as distribution centers, and will directly determine the construction costs. Land prices vary according to local policies and regulations. Should try to choose good policies and low land prices.

The area is also a factor to be considered, because if the distribution center wants to achieve scale development, it needs a certain area to support the storage and circulation of equipment and vehicles. If the area is too large, it will be wasted. If the area is too small, it will not be conducive to later development and expansion.

Topography and terrain take into account traffic and hidden dangers and should be selected slightly higher and flat. Avoid hills, basins and other places.

Infrastructure includes water, electricity, gas and waste disposal.

(2) Natural conditions

Natural conditions mainly include climate, hydrology, and geology. Climate conditions refer to various meteorological indicators such as wind direction, wind force, and precipitation. If you choose to place a large amount of precipitation, it will easily lead to wet, moldy and damaged goods in the open air.

Hydrological conditions refer to the potential impact of flood disasters that need to be taken into account in selected locations. To inquire about past information and understand the local situation, you cannot choose inland areas or flooded areas.

Large-scale transportation vehicles and other facilities and equipment in the distribution center will cause great pressure on the ground or even damage the ground

and cause collapse. Therefore, it is necessary to measure whether the geological conditions can meet the bearing conditions and avoid inappropriate selection of loose soil layer, silt layer, quicksand layer, etc. Geological parcels.

(3) Traffic conditions

Choose convenient transportation construction and distribution center to facilitate multi-modal transportation and daily product turnover, adjacent to transportation hubs (such as national roads, ports, and airports).

(4) Operating conditions

Business conditions including labor, commodity attributes, etc. Adequate supply of labor can ensure the long-term development of the logistics industry; product attributes should also be fully considered when selecting a site. For example, supply-type distribution centers should be located as close as possible to the origin of the materials, and vice versa, sales-type distribution centers. Should be close to consumer groups.

(5) Policies and regulations

Can not violate the national land, environment, construction, transportation and other aspects of the provisions and rigid requirements, and try to find a policy advantage, the construction of a favorable distribution center.

(6) Other conditions

For example: industry competition, environmental protection, etc. Before determining the location, you should find out about the planning of peers in the vicinity, whether there are similar facilities or plans for the construction of logistics centers to avoid vicious competition, but also consider the living environment of nearby residents, do not destroy the ecology and affect people's lives.

3.2.2 Establish Fuzzy Evaluation Model

(1) Establish a fuzzy evaluation index system

The indicator system is divided into three levels taking into account the linkages between the various elements of the evaluation system, as shown in Table 1.

(2) Determine alternatives, evaluation system, and score.

Assume that within a certain area, a logistics center should be selected from n candidate locations. The evaluation system T can be divided into 3 levels.

First: $T = \{T_1, T_2, T_3, T_4, T_5\}$; Second: $T_1 = \{T_{11}, T_{12}, T_{13}, T_{14}\}$, $T2 = \{T_{21}, T_{22}, T_{23}\}$; Third: $T_{14} = \{T_{141}, T_{142}\}$.

The n candidate addresses constitute a decision set $D = \{D_1; D_2; D_3 \ldots D_n\}$. It can be concluded that the scores of each candidate address pass the survey and the corresponding data can be used to obtain a fuzzy comprehensive evaluation table.

Table 1 Fuzzy evaluation system index system level division

First-level indicators		Secondary indicators		Third-level indicators	
Land conditions	T_1	Land price	T_{11}		
		Area	T_{12}		
		Topography	T_{13}		
		Infrastructure	T_{14}	Water, Electricity, Gas	T_{141}
				Waste treatment	T_{142}
Natural conditions	T_2	Climate	T_{21}		
		Hydrology	T_{22}		
		Geological	T_{23}		
Traffic conditions	T_3				
Operating conditions	T_4				
Policies and regulations	T_5				

(3) Determine the weight of each factor set *H*.

Evaluators combine the influence of various factors and give their respective proportions, which are the weights. The weight has a greater influence on the final result of the model. There are many ways to assign weights, and you can choose according to actual needs. This article we use expert consultation method:

$H = \{H_1, H_2, H_3, H_4, \mathrm{H}_5\}$; $H_1 = \{H_{11}, H_{12}, H_{13}, H_{14}\}$; $H_2 = \{H_{21}, H_{22}, H_{23}\}$; $H_{14} = \{H_{141}, H_{142}\}$.

(4) Create an overall evaluation matrix Z and Calculation $Y = H \cdot Z$.

Hierarchical: calculation by weighted average method, $Y_i = \sum(H_i \cdot Z_i)$, calculate the result of Y_{14}, Y_1, Y_2.

General evaluation: First level weight is $H = \{H_1; H_2; H_3; H_4; H_5\}$. Then calculate the value of Y. Sort the calculation results in decreasing order and select the highest value as the best location.

4 Example Analysis

This paper takes Changchun Ouya supermarket as an example. According to the survey, there are 27 Ouya supermarket in Changchun, Puyang store, Jiefang Road store, Linhe store, Hongqi street store, Yueye store, Weixing store, Chuncheng store, Changxin street store, Huxi store, Free road store, Xiaonan street store, Chebai store, Fuan store, Yingbin Road store, Changshen Road store, Liuying Road store, Fusong Road store, Triumphant Road store, Hankou Street store, Suzhou South Street store, Yisheng store, Linyu store, Jingyue second store, Plaza store, Wan Qianyuan store, Dongfeng Street store, Sitong store. Using Baidu map picking coordinate system to

Table 2 The latitude and longitude and coordinates of Ouya supermarket

Number	*X*	*Y*	Number	*X*	*Y*
1	4,861,532	42,442,690	15	4,856,928	42,438,066
2	4,861,257	42,447,297	16	4,868,359	42,444,546
3	4,855,371	42,449,425	17	4,858,540	42,441,844
4	4,860,632	42,444,440	18	4,869,720	42,445,830
5	4,859,609	42,439,652	19	4,864,413	42,445,987
6	4,856,296	42,447,346	20	4,857,681	42,450,981
7	4,862,945	42,441,160	21	4,853,318	42,449,448
8	4,867,176	42,447,318	22	4,853,362	42,443,614
9	4,859,807	42,443,382	23	4,852,510	42,455,625
10	4,859,731	42,448,390	24	4,863,771	42,445,615
11	4,868,741	42,447,815	25	4,861,193	42,447,849
12	4,859,723	42,441,153	26	4,855,745	42,431,993
13	4,862,279	42,450,880	27	4,863,323	42,450,945
14	4,863,881	42,439,867			

obtain the latitude and longitude of each demand point. Then Calculate the coordinates of each demand point by latitude and longitude and XY coordinate conversion software (Table 2).

4.1 Distribution Center Location with Fuzzy Evaluation Method

(1) Solve the coordinates of the center of gravity of each demand point, select the initial location L_0. After investigation, we obtain the average demand rate of demand points and other information. Bring the original horizontal ordinate and demand data into the following formula. Calculate the initial coordinate points $L_0(M_0, N_0)$ with the weighted average method. The result of L_0 is (4,860,529.024, 42,444,451.76).

$$M_0 = \frac{x_1 \times f_1 \times M_1 + x_2 \times f_2 \times M_2 + \cdots + x_n \times f_n \times M_n}{x_1 \times f_1 + x_2 \times f_2 + \cdots + x_n \times f_n} = 4{,}860{,}529.024$$

$$N_0 = \frac{x_1 \times f_1 \times N_1 + x_2 \times f_2 \times N_2 + \cdots + x_n \times f_n \times N_n}{x_1 \times f_1 + x_2 \times f_2 + \cdots + x_n \times f_n} = 42{,}444{,}451.76$$

(2) Calculate the shortest distance and freight cost of each demand point to the initial coordinate point $L_0(M_0, N_0)$. Bring the horizontal and vertical coordinates of

Table 3 Short distance tariff

Distance (km)	≤10	>10
Transportation rate (yuan/t km)	1	0.7

each demand point into Formula (1), then calculate the linear distance from the coordinate point of the initial distribution center. The short-distance transport rate can be obtained through Table 3.

Calculate the initial shipping cost C_0 and the total initial shipping cost from the demand point to the initial delivery center coordinate point by Formulas (2–4). $C_0 = 711$ yuan.

(3) Substitute j_i and find new location L_1. And find the new total distribution cost C_1.

Bring requirements, transportation rates, distances, coordinates, etc. into Eqs. (7) and (8) which are Obtained through investigation. Calculate a possible best distribution center location $L_1(M_1, N_1)$. The coordinate is (4,860,549.11, 42,444,226.04). Calculate the total cost C1, C1 = 707 yuan.

(4) Compare C_0 and C_1 to determine whether to end the calculation.

If $C_0 > C_1$, the initial distribution center location is not the best location. We need to bring point (M_1, N_1) into the above formula to recalculate. If we make sure $C_{k+1} \geq C_k$ when (M_{k+1}, N_{k+1}), then (M_k, N_k) is the best location.

The iterative solution to the optimal position is shown in Table 4.

From Table 3, it can be seen that the difference between continuing iteration costs gradually becomes smaller, so the iteration ends. When distribution center location is L_{10} (4,860,211.74, 42,443,023.83), total cost is minimum. This point is the optimal location point. Convert to longitude and latitude coordinates (125.291133,

Table 4 Iterative process

Iterations k	Mk	Nk	Ck
0	4,860,529.02	42,444,451.76	710.6
1	4,860,546.80	42,444,220.17	706.6
2	4,860,485.10	42,444,001.33	703.6
3	4,860,413.82	42,443,759.72	700.6
4	4,860,352.89	42,443,546.29	698.7
5	4,860,304.74	42,443,379.72	697.5
6	4,860,269.27	42,443,257.41	696.9
7	4,860,245.28	42,443,169.29	696.5
8	4,860,229.63	42,443,105.42	696.4
9	4,860,219.14	42,443,058.57	696.3
10	4,860,211.74	42,443,023.83	696.2
11	4,860,206.31	42,442,997.86	696.2

43.874534). Based on the map, the location is located in the inner part of the Huaqiao village which is near Huxi west road and kaiyun street.

(5) Combined with other conditions for comprehensive analysis.

Due to the Huaqiao village is residential area. The population is dense and there is no land available. So take this point as the center of the circle to find the right place. The north is a residential area and will not be considered. In the east, it is also a residential area, and Hongqi Street is a commercial street that is more expensive and will not be considered. There are also some companies and communities south of Huxi Road. The buildings are also dense and it is not appropriate to choose this place. There are some open spaces at the intersection of the west of Huxi Bridge and Huxi Road. Named as standby center 1, standby center 2, and standby center 3 respectively.

4.2 Distribution Center Location with Fuzzy Evaluation Method

The preliminary location of the distribution center of Ouya supermarket has been obtained and is located in Huaqiao village area with centroid method. We should choose one of the three candidate locations to build a logistics center.

(1) Determine the evaluation system and alternative sites, and obtain a fuzzy comprehensive evaluation table.

The evaluation system T can be divided into 3 levels.

First: $T = \{T_1, T_2, T_3, T_4, T_5\}$; Second: $T_1 = \{T_{11}, T_{12}, T_{13}, T_{14}\}$, $T2 = \{T_{21}, T_{22}, T_{23}\}$; Third: $T_{14} = \{T_{141}, T_{142}\}$. The three candidate addresses constitute a decision set $D = \{D_1; D_2; D_3\}$. We get he scores of each candidate address through the survey and use the corresponding data to obtain a fuzzy comprehensive evaluation table as Table 5.

(2) Determine the weight of each factor set *H*.

Using expert consultation to determine weights as follows:

$H = \{0.3; 0.2; 0.5; 0.2; 0.1\}$; $H_1 = \{0.4; 0.3; 0.2; 0.2\}$; $H_2 = \{0.2; 0.1;0.3\}$; $H_{14} = \{0.3; 0.2\}$.

(3) Fuzzy evaluation method

Create a total evaluation matrix *Z*, Calculate $Y = H \times Z$. Hierarchical:

$T_{14} = \{T_{141}, T_{142}\}$, Weight: $H_{14} = \{0.3; 0.2\}$. The constituent single factor matrix: $Z_{14} = \begin{Bmatrix} 3 & 3 & 3 \\ 2 & 1 & 1 \end{Bmatrix}$, $Y_i = \sum(H_i \cdot Z_i)$, $Y_{14} = \sum(H_{14} \cdot Z_{14}) = \{1.3, 1.1, 1.1\}$.

Table 5 Reserve survey table

Alternative address						NO. 1	NO. 2	NO. 3
Evaluation conditions								
Land conditions	T_1	Land price			T_{11}	2	2	2
		Area			T_{12}	4	3	2
		Topography			T_{13}	2	3	3
		Infrastructure	T_{14}	Water, Electricity, Gas	T_{141}	3	3	3
				Waste treatment	T_{142}	2	1	1
Natural conditions	T_2	Climate			T_{21}	3	3	3
		Hydrology			T_{22}	1	2	1
		Geological			T_{23}	3	4	4
Traffic conditions					T_3	3	3	3
Operating conditions					T_4	3	2	2
Policies and regulations					T_5	1	1	1

Same reasoning, $\mathrm{Y}_1 = \sum(\mathrm{H1} \cdot \mathrm{Z1})$, $\mathrm{H}_1 = \{0.4; 0.3; 0.2; 0.2\}$, $Z_1 = \begin{Bmatrix} 2 & 2 & 2 \\ 4 & 3 & 2 \\ 2 & 3 & 3 \\ 1.3 & 1.1 & 1.1 \end{Bmatrix}$,

$Y_1 = \{2.66, 2.52, 2.22\}$. $Y_2 = \sum(H_2 \cdot \mathrm{Z}_2)$, $H_2 = \{0.2; 0.1; 0.3\}$, $Z_2 = \begin{Bmatrix} 3 & 3 & 3 \\ 1 & 2 & 1 \\ 3 & 4 & 4 \end{Bmatrix}$,

$\mathrm{Y}_2 = \{1.6, 2.0, 1.9\}$.

General evaluation:

$$\mathrm{Y} = \sum(\mathrm{T} \cdot \mathrm{Z}),\ \mathrm{H} = \{0.3;\ 0.2;\ 0.5;\ 0.2;\ 0.1\},\ Z = \begin{Bmatrix} 2.66 & 2.52 & 2.22 \\ 1.6 & 2.0 & 1.9 \\ 3 & 3 & 3 \\ 3 & 2 & 2 \\ 1 & 1 & 1 \end{Bmatrix},\ \mathrm{Y} =$$

{3.318, 3.156, 3.046}.

(4) Conclusion

The comprehensive evaluation results of alternative sites are ranked as: No. 1 > No. 2 > No. 3. The land No. 1 represented by the highest value is selected as the best distribution center location, the longitude and latitude coordinates are (125.284809, 43.876731), and the *XY* coordinates are (4,860,460, 42,442,518).

5 Conclusion

This study is base on the survey of 27 Ouya supermarkets in Changchun to conduct an initial site selection of the distribution center with centroid method. It explains the factors affecting the site selection, and uses fuzzy evaluation methods to focus on the three alternatives near the initial site selection. Finally receive the location plan of distribution center of Ouya supermarkets in Changchun with qualitative and quantitative analysis. Due to the demand of supermarket is dynamic, but also influenced by many other uncertain factors, how to design a can adapt to change, constantly adjust the scientific location of distribution center will be one of the focus of future research.

References

1. Drezner Z, Klamroth K, Schobel A, Go W (2002) The Weber problem. In: Drezner Z, Hamacher H (eds) Facility location: application and theory. Springer, London, 1–36
2. Hoover EM (1957) Location in theory and leather industries. Harvard University Press, Cambridge, MA
3. Hakimi SL (1964) Optimum location of switching centers and the absolute centers and medians of a graph. Oper Res 12:450–459
4. Baumol W, Wolfe PA (1958) Warehouse location problem. Oper Res 6(2):252–263
5. Alan TM, Ross AG (1997) Capacitated service and regional constraints in location-allocation modeling. Locat Sci 5(2):103–118
6. Kuehn A, Hamburger M (1963) A heuristic for location warehouses. Manag Sci (6):643–666
7. Spohrer GA, Kmak TR (1984) Qualitative analysis used in evaluating alternative-plant location scenarios. Ind Eng (8):52–56
8. Gao X, Li Z (1994) Logistics center location model and a heuristic algorithm. Oper Res Manag Sci 3(3–4):56–62
9. Li Q, Yuan Q, Wen D (1999) Location model under optimal inventory strategy. Syst Eng 17(6):7–11
10. Ding H, Li D (2004) Research on location method of city logistics distribution center. J Huazhong Univ Sci Technol (Urban Science Edition) 01:52–56
11. Yang B (2003) The problem of logistics distribution center location with multi-species random mathematic model. Chin J Manag Sci 11(2):35–39
12. Wang H, Sun S (2011) Discussion on location of logistics center based on gravity method. J Commer Econ 35:39–40
13. Wang F, Lin W (2014) Research on location selection of 3PL distribution center based on improved center of gravity-causal analysis. Logist Technol 33(3):185–187
14. Zhao Y, Zhang Y (2008) Location optimization and application of chain supermarket free distribution center based on center of gravity method. Logist Technol 27(10):116–117
15. Qin Z (2011) Analysis of distribution center location of third-party logistics enterprises based on cluster analysis and gravity method—taking Ganyun logistics as an example. Logist Technol 30(8):101–103

A Bayesian Recognition Method for Highway Ambiguity Path Identification Based on Digraph

Xu-jin Yu, Jun Xu and Fan Zhang

Abstract With the rapid development of the highway, the layout of road network is gradually transformed from the tree structure to the reticular structure. In order to develop a fair charging strategy, it is necessary to split the route of vehicles accurately. In this paper, a directed graph based method for ambiguous path of highway identification with Bayesian recognition algorithm is studied. Firstly, the highway network should be modeled into a digraph structure, and the 5.8G identification point or video image identification point would be mapped to the directed graph. Then a Bayesian recognition algorithm would be used to determine the actual path of the vehicle. The algorithm uses a two-stage structure, which could provide a rapid calculation at the exit lane. The method has been applied in the highway of Jiangxi Province and has achieved valuable results.

Keywords Intelligent transportation · Ambiguous path identification · Directed graph · Bayesian

1 Introduction

Recently, with the rapid development of highway construction in China, the layout of highway network is gradually transformed from the tree structure to the net structure. There are several possible paths between the entrances and exits. Due to the diversification of highway construction investors, the research on vehicle ambiguity path recognition involves not only how to calculate the toll rates, but also how to split tolls to each investment entity. Therefore, the solution of ambiguous path identification is the key to make sure the toll can be accurate, reasonable and fair distribution [1–4].

X. Yu · J. Xu
Jiangxi Provincial Highway Network Management Center, Nanchang 330036, China

F. Zhang (✉)
Research Institute of Highway, MOT, Beijing 100088, China
e-mail: zhangfan@itsc.cn

Beijing GOTEC ITS Technology Co., Ltd., Beijing 100088, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_3

The shortest path method is widely used in highway toll rate calculation, through the road network analysis of the shortest path from the entrance to the exit, to achieve the calculation and distribution of the toll collection. Since the complexation of the highway network is increasing, this method will inevitably lead to more and more ambiguity in toll calculation and directly impact the interests of each highway construction investor [3, 4].

In order to identify the actual path of the vehicle, it is necessary to set up the identification station in the road network to record the path of the vehicle [5, 6]. The identification stations can use a variety of identification technology including 5.8G DSRC [7] (Dedicated Short Range Communications) technology, 433M radio frequency technology [8, 9], video license plate recognition technology [10, 11], communication operator base station identification and so on. Those technologies record the vehicles passing-through identification site information into the OBU (On Board Unit) or CPC (Composite Pass Card) carried by vehicles, or send them directly to toll stations. When the vehicle reaches the exit lane, the path of the vehicle will be analyzed by reading the information, and the charging and toll distribution are also carried out at the same time. The shortest path between the identification points is usually used in the analysis process.

This generalized ambiguity path recognition method has two major problems. Firstly, the path identification system is unreliable under some circumstance, which leads to a missed mark or a false mark. For instance, Video license plate recognition technology is extremely easy to be influenced by weather conditions, and other problem such as license plate defaced, plate covered and plate faked, due that, the recognition rate can only reach about 90%. 433M radio frequency technology has a long radiation distance, which make it easy to misidentify vehicle passed by the station. The complexity of the signal from communication operator base station will also produce more missing or mislabeled label. Although 5.8G DSRC technology identification accuracy can reach more than 99%, due to card aging, identification antenna failure, external interference and other reasons, there is also a certain leakage probability. Secondly, the traffic flow of highway exit lane is very heavy, especially the ETC vehicle will pass through the exit in a very short time, which need to use a fast ambiguity path analysis method.

In this paper, a directed graph based method for ambiguous path of highway identification with Bayesian recognition algorithm is proposed. The complex highway network is modeled as a digraph, and the identify points are mapped into the digraph. Then the Bayesian recognition algorithm is used to analyze the actual path of the vehicle based on the vehicle identification station information. In the process of realization, the proposed algorithm uses a two-stage structure, road network and identification points are preprocessed in the management center, and a rapid marching calculation is used to make sure the vehicle pass the exit lane quickly.

The organization of this paper is as follows, the second section of this paper introduces the process of the digraph based method for ambiguous path of identification with Bayesian recognition algorithm. Then the experimental results of the proposed algorithm is presented in the third section. At the end, the conclusion of this algorithm will be shown.

2 Bayesian Recognition Method for Highway Ambiguity Path Identification Based on Digraph

2.1 Highway Network Modeling and Identification Station Mapping Based on Digraph

Highway network topology contains many elements, including interchanges, entrances and exits, and vehicle can also make a U-turn before the toll station. There are several research studies have been done for the highway network topology modeling with Geographic Information System [12, 13]. In these studies, the intersection, interchange, entrance and exit of the network are simplified into a node, each node is connected by the road section and establishes a graph structure. Chen et al. further considered the ramp information of interchange and added directional information to the nodes [14]. However, this is still considered as the undirected graph model, and the up-down directions of the roads were taken as the same edge of the road section.

In order to accurately describe the driving path of the vehicle, the vehicle path must be modeled, including distinguishing the direction of the road, describing the ramp relationship at the interchange entrance and exit, and the U-turn position of the road. This paper uses a directed graph to model the road network. For the convenience of description, the following definitions are given:

Definition 1 (*link a*) Link a is a road segment with the determined direction, which has only a single entry point and exit point, separately as $v_{start}(a)$ and $v_{end}(a)$. Vehicles in the road driving paths and directions are completely determined, the length of the link is defined as $length(a)$.

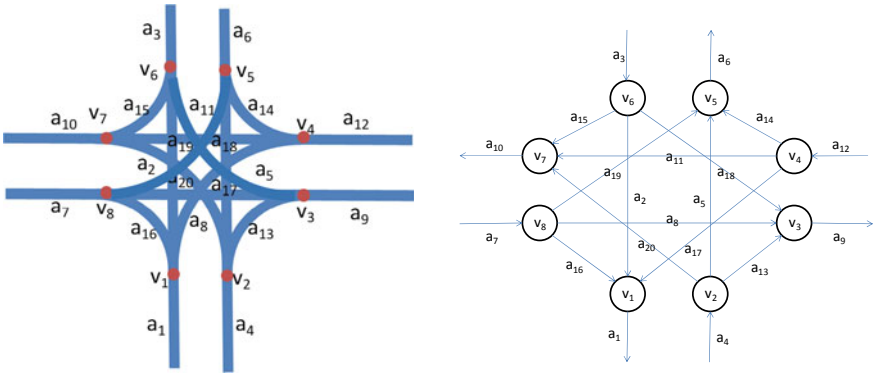
Definition 2 (*Node v*) Node is defined as the point of highway where a link is intersected by another link, such as the entrance, exit, on-ramp, off-ramp and U-turn point. All entrance and exit of the link are defined as node. V_{entry} is denoted the set of entrance nodes on the whole network and V_{exit} is donated the set of the exit nodes on the whole network.

According to the above definition, the highway network can be represented as a digraph $\overrightarrow{G}(V, A)$, where V is the set of the nodes in the network and A is the set of the links in the network. The following common highway interchanges in Figs. 1a and 2a could be modeled as the digraph in Figs. 1b. and 2b.

Interchange of highway and its digraph is shown in Fig. 1.

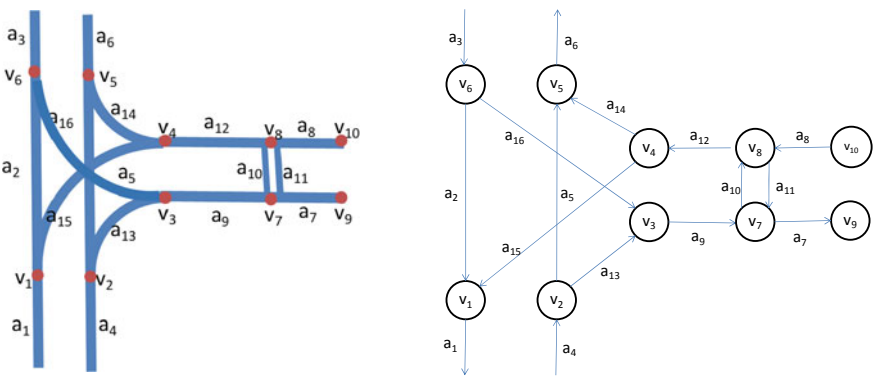
Entrance and exit of highway and its digraph is shown in Fig. 2. The length of a_7, a_8, a_{10}, a_{11} in Fig. 2b are defined as zero.

When a vehicle passes through a road section, according to the order of the driving path, it may be marked by one or more identification stations, such as being marked by a 5.8G multi-path antenna tag or by a HD license plate recognition identification station. There are some studies about optimizing location methods of the identification station, in the study, the path of the vehicle can be analyzed by using fewer



(a) Interchange of highway (b) digraph of interchange

Fig. 1 The interchange of highway and the digraph



(a) the entrance and exit of highway

(b) directed graph of entrance and exit

Fig. 2 Entrance and exit of highway and the digraph

identification stations in the road network. Due to the possibility of missing labels for each type of marking technology, there is also the possibility of marking result while a vehicle passes through a road section. Moreover, according to the marking technology of 5.8G DSRC and 433M, both sides of the road are generally covered by the antenna signals at the same time, not only are the passing vehicles on this road section identified, but also the vehicles in the opposite direction are also identified.

For all identification stations $m_1, m_2 \ldots m_k$ on link a, $p(a, i)$, $i = 1 \ldots k$ means the probability that station i will be marked. An experienced value of $p(a, i)$ could be defined according the technical type of the identification station. For instance, the value of 5.8G technology would be 99.5% and the value of video license plate

recognition technology would be 90%. This value also could be obtained from the learning process of historical data.

When the vehicle passing link a, an identifying string $s = m_{i_1} m_{i_2} \ldots m_{i_n}$, $1 \leq i_1 < i_2 < \cdots < i_n \leq k$ will be created. The probability of producing string S is in below:

$$p(s|a) = \prod_{j=1\ldots n} p(a, i_j) \times \prod_{l \in [1,k], l \neq i_j} (1 - p(a, l))$$

2.2 Building Path Tree of Every Exit

Definition 3 (*vehicle path P*) Vehicle path P refers to the road consists of links and nodes connected by one or more orders which starts at the entrance and ending at the exit of a highway. The entrance and the exit of the path are separately denoted as $v_{start}(P)$ and $v_{\text{end}}(P)$, $v_{start}(P) \in V\_Entry$, $v_{\text{end}}(P) \in V\_Exit$. Road nodes of the vehicle path are denoted as $v_0(P) = v_{start}(P) = v_1(P)$, $v_2(P) \ldots v_n(P) = v_{\text{end}}(P)$. The links are denoted as $a_1(P)$, $a_2(P) \ldots a_{\text{n}}(P)$, $\forall i = 1 \ldots n$, $v_{i-1}(P) = v_{start}(a_{\text{i}}(P))$, $v_i(P) = v_{end}(a_{\text{i}}(P))$.

In order to obtain all the vehicle paths of each freeway exit, it is necessary to build path tree for each exit point using depth first or breadth first algorithm along the road direction after establishing the topology of the road. When building a path tree, the loop path should be considered because of the ring structure or lost of vehicles for a variety of reasons. The maximum number of rings is denoted as $N\_MaxLoop$, and the number on the same ring cannot be more than $N\_MaxLoop$.

The specific path tree algorithm steps are as follows:

(1) Building a tree $Tree(v_0)$ for each exit node $v_0 \in$ V_$Exit$ which is considered as root node.
(2) Performing step (3) for each leaf node of the tree $Tree(v_0)$.
(3) If v_1 is not the freeway entrance $v_1 \notin V$_Entry, finding each link a with this node as the exit in digraph $\overrightarrow{G} = (V, A)$. Then performing (4), (5), (6) for link a.
(4) Finding $v_{\text{start}}(a)$ for each link a.
(5) Finding the number $N(v_{start}(a))$ of $v_{\text{start}}(a)$ from v_1 to root node v_0 in $Tree(v_0)$.
(6) If $N(v_{start}(a)) < N\_MaxLoop$, then $v_{\text{start}}(a)$ is denoted as a sub-node of v_1, and link a is denoted as connected relationship between $v_{\text{start}}(a)$ and v_1.
(7) Return to (2) until all the leaf nodes of $Tree(v_0)$ are entrances of freeway.

There is a vehicle path between each leaf node and root node in $Tree(v_0)$, and there are many vehicle paths between the same entrance and exit. Figure 3 shows a sample of path tree which has three paths between entrance v_7 and exit point v_0.

All vehicle paths for the same entrance and exit are denoted as $PathSet(v_1, v_0) = \{P|v_{start}(P) = v_1, v_{end}(P) = v_0\}$.

There is a probability $p(P)$ for choosing a path P in $PathSet(v_1, v_0)$ between entrance v_1 and exit v_0. This probability can be estimated by the distance comparison between different paths, the charge comparison, the congestion degree, etc. It can also be obtained according to the vehicle historical path data.

2.3 Building Probabilistic Model of Identification String Using Bayesian Recognition Algorithm

The vehicles pass through each link $(a_1(P), a_2(P) \ldots a_n(P))$ contained of the vehicle path and generate the corresponding identification string. They are denoted as $S = s_1 s_2 \ldots s_n$. The probability of generating the identification string can be obtained by the probability calculation of the corresponding sub-identification string of each link on the path. The probability is as follows:

$$p(S|P) = \prod_{i=1\ldots n} p(s_i|a_i(P)).$$

Under the same entrance and exit conditions, the same identification string may be generated when vehicles pass through different vehicle paths. For example, in Fig. 3, if there is no identification station, or a missing identification in link a_3, a_4, a_8, then path $P_1 : v_5 \rightarrow v_3 \rightarrow v_1 \rightarrow v_0$ will generate the same identification string with path $P_2 : v_5 \rightarrow v_1 \rightarrow v_0$.

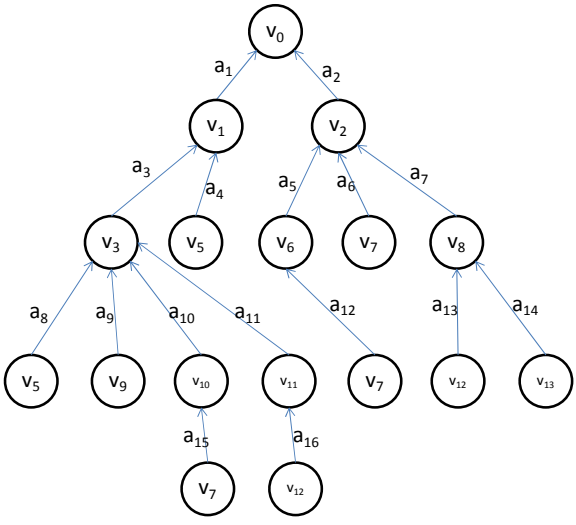


Fig. 3 Vehicle path tree

The probability of the vehicle passing a certain path using Bayesian recognition algorithm for the result of the identification string of the vehicle is as follows:

$$p(P|S) = \frac{p(S|P)p(P)}{\sum_{P_i \in PathSet(v_1, v_0)} p(S|P_i)p(P_i)}$$

When the vehicle arrives at the exit, the export management system reads the entrance and identification stations information along the way from the car unit or the road network information center. The vehicle path with maximum probability can be obtained based on the identification string as $P_{\max}(S)$, $\forall P \in PathSet(v_1, v_0)$, $p(P|S) \leq p(P_{\max}|S)$.

This path can be used as the vehicle's final path. On the basis of calculating determine the final path information, the expense, calculated on the basis of the piecewise finally merger way, according to the models, the number of shaft, the weight, and each segment mileage to calculate the final billing amount.

3 Algorithm Implementation

In the freeway, the algorithm needs to consider the rapid calculation of the exit, especially the ETC vehicle, and generally requires the cost analysis to be completed within 500 ms. This algorithm can be divided into two parts: the pretreatment process of the provincial center and the rapid calculation process of the path exit.

The pretreatment process of the provincial center:

(1) Building digraph $\overrightarrow{G} = (V, A)$.
(2) Building $Tree(v_0)$ for each exit node v_0.
(3) Building set of path from v_1 to v_0 for each entrance node v_1 as $PathSet(v_1, v_0)$.
(4) Analyzing all the identification string sets.
(5) Building the relationship between S and $P_{\max}(S)$.
(6) Sending this relationship $(v_1, S) \rightarrow P_{\max}(S)$ to the exit.

The rapid calculation process of the path exit:

(1) Obtaining entrance node and identification string of the vehicle.
(2) Obtaining $P_{\max}(S)$ of the vehicle according to the corresponding relational table.
(3) Calculating the vehicle cost according to $P_{\max}(S)$

The computation is simple because the exit is just a simple table lookup process.

Table 1 Comparison of vehicle path recognition accuracy of different algorithms

Algorithms	The shortest path algorithm based on the entrance and exit	The shortest path algorithm based on identification points	Digraph algorithm
Number of accurate identification	10,181	10,437	10,662
Accuracy of identification (%)	95.3	97.7	99.8
Average computation time (ms)	1.5	7.8	2.1

4 Experiment Results

To test the effect of the current algorithm, a monthly data experiment was conducted in 89 experimental vehicles (operating vehicles) in Jiangxi province. In total, the path information of 10,683 times is processed by the algorithm of this paper and the shortest path algorithm, and compared with the actual path. The computing platform adopted is strong E5 3G CPU X2, 16GMEM. The results are shown in Table 1.

Through Table 1 analysis of experimental results shows that the ambiguity path identification algorithm based on digraph has greatly improvement than shortest path algorithm, and no path can't be identified. The algorithm can be improved in path query time.

5 Conclusion

In this paper, the freeway network is modeled and mapped to a digraph and using Bayesian algorithm to analyze the vehicle path. The calculation speed of the algorithm is improved to ensure the fast passing of the vehicle through the two steps of network center and toll station. This algorithm has been applied to the expressway network in Jiangxi province, and it has good adaptability to the omission and error, and the processing time can fully meet the demand of freeway exit. The next step is to further expand the algorithm and support the new business such as the differential rate of the freeway. In addition, the time and path distance information of the vehicle is used for further reduce the error of vehicle path identification. The algorithm can also be used for alarm processing of abnormal matches, and help highway management to prevent vehicle rewind and other escape fees.

Acknowledgements The authors would like to thank the project of "The key technology of the data analysis for the fresh agricultural products vehicles credit audit".

References

1. Song Z, Zhao X (2009) Research on precise charge and split for expressway vehicle tolls. Highway Eng 1:147–150
2. Luo T, Wu C (2016) The research of highway toll ambiguity clearing rules. J Shanghai Sci Res Inst Shipp 2:56–60
3. Ren Y (2014) Research on the problem and countermeasures of the ambiguity path of BOT project. Transp Inf Ind 6:74–78
4. Wang Q (2010) Research on the toll clearing method of complex highway network. Wuhan University of Technology
5. Wu R (2015) Application analysis of ambiguous route identification techniques. Qinghai Jiaotong Keji (6):13–14, 35
6. Sun K (2013) Research and realization of highway multi-path identification technology. Zhengzhou University
7. Lang J, Xu J (2015) Research on the ambiguity of highway network based on 5.8G ETC. China Highway (z1): 225–228
8. Li Y (2014) Research on the precise splitting method of highway toll based on RF-SIM card. Changsha University of Science and Technology
9. Zhou J, Peng J (2012) Application analysis of RFID technology on experimental scheme of Guangdong ambiguous path identification. Guangdong Highway Commun 4:56–60
10. Gu Y (2014) Research on license plate recognition system. Ningxia University
11. Li X (2013) Design and implementation of integrated application system based on video detection technology. Hebei University of Science and Technology
12. Sun Y, Chen S, Huang H (2003) Adaptive optimal route selection based on gray evaluation theory. China J Highway Transp 4:87–90
13. Jin K, Li C, Qin Q (2006) Study on shortest path search method based on ant algorithm. J Highway Transp Res Dev (3):128–130, 134
14. Chen Y, Chen S (2010) Complexity topology of expressway network with information of interchange ramps. J Tongji Univ (Natural Science) 2:230–237

Time-Varying Characteristics and Forecasting Model of Parking Berth Demand in Urban Residential Areas

Jun Chen, Yi-fan Yue, Jingheng Zheng and Dong-ping Li

Abstract In order to improve the micro analysis and prediction of real-time forecasting method of dynamic parking demand, we selected three typical residential areas in Yangzhou City as an example to analyze the time-varying characteristics of motor vehicles' arrival and departure. Considering the obvious difference between the arrival and departure characteristics of motor vehicle in residential areas on weekdays and weekends, the different time series models were used to forecast the berth occupancy of three residential areas on weekdays and weekends. Due to the higher proportion of commute travel on weekdays and the higher proportion of flexible travel on weekends, the variation tendency of berth occupancy on weekends is not as stable as that on weekdays. The result shows that the prediction accuracy of real-time numbers of berth on weekdays is usually higher than that on weekends. On weekdays, the berth occupancy rate of three residential areas is regular, which can be forecasted by ARIMA (Autoregressive Integrated Moving Average) model, and can reach more than 98% of the prediction accuracy. Oppositely, the weekends' time-varying regularity of berth occupancy is not obvious, thus using ARMA (Autoregressive Moving Average) model, and the accuracy can reach over 95%. Overall, time series model has good adaptability to the residential area, and the higher accuracy can be achieved by selecting the appropriate model.

Keywords Urban residential areas · Time-varying characteristics · Parking demand forecast · Model

J. Chen (✉)
Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Nanjing 210096, China
e-mail: chenjun@seu.edu.cn

Y. Yue
School of Transportation, Southeast University, Nanjing 210096, China

J. Zheng
Suzhou City Transportation Management Office, Suzhou 215008, China

D. Li
Shanghai Urban Construction Design and Research Institute (Group) CO., LTD, Shanghai 200125, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_4

1 Introduction

With the expansion of city scale and the increase of the citizen's income of residents, the proportion of travelling by private cars is rising. However, urban parking resources are limited by the urban space and land prices and other factors, the speed of construction of urban parking resources in China obviously lags behind the speed of city motorized, which makes the contradiction of parking supply and demand seriously, especially in urban residential area. The parking of old residential areas near the city center is particularly serious. Old residential areas particularly suffer from serious parking contradiction problem, and the parking behavior there appears obvious time-varying characteristics. How to effectively alleviate the contradiction between urban parking and optimize the allocation of parking resources based on time-varying characteristics of parking berth demand in residential areas, has become an urgent and necessary need. Moreover, it is of great theoretical and practical value.

As for forecasting model of parking demand, Levinson [1] based on land use and employment data, used a similar gravity model to predict parking demand in urban peak times. Vlahogianni [2] predicted from two different angles: the average free time and berth occupancy rate at preset time period, the study found that the Weibull function model can better fit the average free time, and the genetic optimization of multilayer perception can accurately predict parking demand in the future per 30 min. Han [3] introduced the concept of regional development factors and by using the survey data, has carried on the forecast demand for parking, modified the generation rate model by giving full consideration on different factors of the economy, society, transportation and level of parking demand. Xue [4] based on the method of parking generation rate, considering the sharing of parking resources between different land types and the discount of shared parking spaces, built a shared berth prediction model. Qin [5] analyzed the parking demand characteristics of the typical mixed land of Beijing trade center through the parking demand survey, and established a model of parking demand superposition model and parking demand sharing model based on single purpose building. Guan [6] considered factors of price and parking service level influence on parking demand, and respectively divided parking demand into day and night to make predictions, built a prediction model of parking demand-supply based on the generation rate model.

However, few research covered micro forecast of dynamic parking berth demand considering the time-varying regularity, the focus is mainly on the macroscopic forecasting method of parking demand, or just put forward a theoretical method of microscopic prediction without data support.

In order to solve these problems above, this paper takes typical residential areas as research object, will focus on the real-time change regularity of parking berth from the micro point of view. Then forecast the real-time berth occupancy in residential areas by the method of time series, and verify the time-varying characteristics of berth occupancy in residential areas on weekdays and weekends.

2 Analysis on Time-Varying Characteristics of Parking Berth Demand in Urban Residential Areas

2.1 Survey Design

In order to grasp the real-time occupancy of parking berth in urban typical residential areas and obtain the distribution data of vehicles' arrival and departure, considering the three factors of scale, location and built year of residential area, this paper selected three representative typical residential areas in the central area of Yangzhou City (Yangzhou middle school dormitory, Zhihuai village and Shuangqiao village) as survey respondents. By using video recording method and observing the entrance and exit of the three areas, the data of arrival and departure of motor vehicles in three residential areas were real-timely recorded and quantitatively analyzed at 5-min intervals for nine consecutive days from December 21, 2013 to December 29, 2013 (including five working days and four rest days).

Assume that the arrival number of motor vehicles at each time is A_n, the departure number is L_n, and the number of real-time berth occupancy in residential area is P_n. According to the survey method mentioned above, the value of P_n at each time can be calculated:

$$P_{n+1} = P_n + A_n - L_n \tag{1}$$

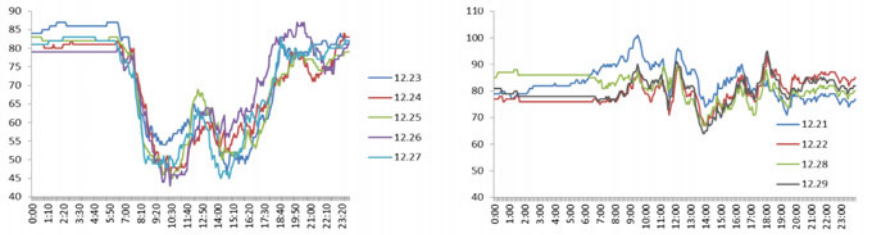
where P_n denotes the real-time number of berth occupancy in residential area on time n, A_n denotes the arrival number of motor vehicles in residential area on time n, L_n denotes the departure number of motor vehicles in residential area on time n, and P_{n+1} denotes the number of berth occupancy in residential area on time n + 1.

2.2 Analysis on Occupancy Characteristics of Parking Berth in Residential Areas

According to the initial number of berth occupancy and the obtained data of arrivals and departures, the initial time for recording is at 9:00 am on December 23, 2013, we can calculated the real-time number of berth occupancy of the three residential areas at each time period by Formula (1). Take Area 1 as an example, the berth occupancy curve on weekdays and weekends are respectively shown as follows in Fig. 1.

As can be seen from Fig. 1, the curves of parking berth condition in residential area show different occupancy characteristics on weekdays and weekends. Based on parking data of Area 1, the time-varying occupancy characteristics of parking berth are summarized as follows:

(1) The time-varying occupancy condition of parking berth in residential area 1 on weekdays is of significant regularity, and the peak period usually starts from



(a) Berth occupancy on weekdays (b) Berth occupancy on weekends

Fig. 1 The curve of berth occupancy in Area 1

around 7:00 pm and lasts until 8:00 am the next day. Then the number of berth occupancy is drastically reduced as the residents go to work or school. At around 11:00 am to 2:00 pm, a small peak appears in the berth occupancy, since some residents may take siesta at noon.

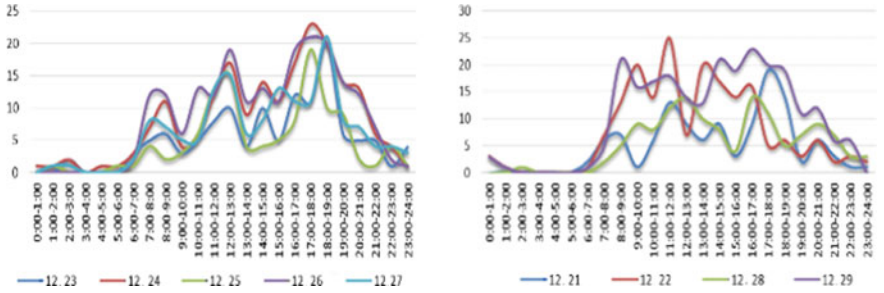
(2) Compared with the weekdays, the number of berth occupancy at Area 1 has no obvious regularity on weekends. The occupancy of parking berth on weekends is constantly changing, there are more random fluctuations from 8:00 am to 8:00 pm throughout the day.

2.3 Analysis on Arrival and Departure Characteristics of Motor Vehicles in Residential Areas

Because of commute travel, there will be relatively large differences in the travel and parking behavior of car users in residential area on weekdays and weekends. In order to be able to more accurately understand the time-varying characteristics of parking behavior in residential area, this paper will respectively discuss the arrival and departure regularity of motor vehicles on weekdays and weekends. Based on the previous observation for nine consecutive days of surveillance video, the numbers of cars entrance and exit the residential area were respectively recorded at every five-minute interval, then make the recorded data be organized into hourly units at each whole point of time. Take Area 1 as an example, the statistical results are as shown in Figs. 2 and 3.

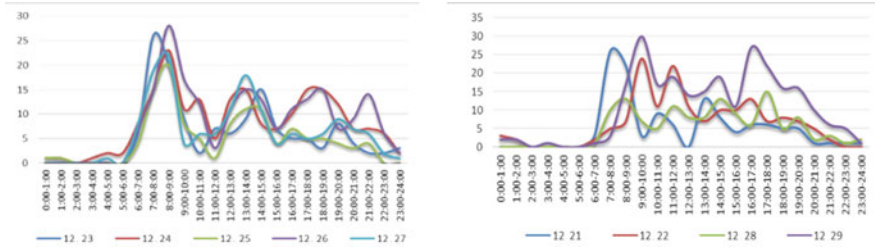
(1) Arrival characteristics

1. It can be seen from the Fig. 2, the motor vehicles arrival curves in residential area on different weekdays are approximately similar, and there is certain regularity in the tendency of the motor vehicles arrival in residential area: The arrival curve on weekdays shows bimodal distribution, and there will be two peaks of motor vehicles arrival throughout the whole weekday, while the change of vehicles arrival during the rest time is relatively stable. Where the



(a) Vehicles arrival on weekdays (b) Vehicles arrival on weekends

Fig. 2 The curve of vehicles arrival in Area 1



(a) Vehicles departure on weekdays (b) Vehicles departure on weekends

Fig. 3 The curve of vehicles departure in Area 1

first peak ranged from 10:00 am to 1:00 pm, and the second peak appeared between 4:00 pm and 7:00 pm, moreover, the second peak of vehicles arrival in the afternoon on weekdays was more pronounced than at noon. The observed data were consistent with daily travel habits, which the formation of the first arrival peak was related to some residents' behavior of returning home after work at noon, the second is due to the afternoon commute arrival.

2. Compared with the motor vehicles arrival on weekdays, the arrival characteristics on weekends are not obviously regular. And there is no significantly concentrated period of vehicles arrival in the daytime of weekends, except few vehicles arrive before 7:00 am, the vehicles arrival distribution during the rest of time on weekends is of large randomness and shows no evident arrival peak, since motor vehicles can be observed entering into the residential area at all time periods.

(2) Departure characteristics:

1. According to the results shown in Fig. 3, the departure characteristics of motor vehicles in residential area on weekends also have certain regularity, which show a trend of multimodal distribution and will appear three departure peaks throughout the day on weekends. The first and most obvious

departure peak occurred between 7:00 am and 9:00 am in the morning, the second peak started from around 12:00 am to 2:00 pm, and the third one appeared between 5:00 pm to 8:00 pm. On the whole, the variation trend of motor vehicles departure in residential area on weekends is basically the same as the daily travel habits of the residents. The first peak is owning to commute travel in the morning while the second commute peak is at noon, the third one is for elastic travel in the evening.

2. The trend of motor vehicles departure in residential area on weekends can not be obviously noticed, besides, there is no obvious peak value reflecting in the vehicles departure, which is also related to daily travel habits. This is because the travel purpose of most residents on weekends varies from that on weekdays, there is less rigid demand for commuting on weekends, mainly for random elastic travel needs.

3 Forecasting Model of Parking Berth Demand in Urban Residential Areas

Time series models have been the focus of considerable research and development in recent years in many disciplines, including transportation. This interest stems from the insights that are gained when observing and analyzing the behavior of a variable over time; the principle is to predict the future estimates of the variables based on a series of past values at regular time intervals by observing and analyzing the continuity and regularity of the variables over time. A time series is a sequence of observations arranged by their time of outcome, such as the number of vehicles arriving in the residential area for a continuous period of time. The typical characteristic of the time series is the interdependence between adjacent observations.

In time series, there are many unknown factors in the independent variables used to predict the future value of the dependent variable, which can not accurately predict the sequence with a deterministic model. Thus, the mathematical model commonly used in time series is stochastic model, which is often used to achieve optimal prediction and control by calculating the probability that a future value falls within a particular interval.

The stability of time series is defined as follows: If the mean and variance of a random time series is constant over time, and the covariance of any two periods only depends on the distance of the two moments or lag rather than the actual time of the covariance, then it is called stable time series. There are three common types of time series models developed in most research: the autoregressive models (AR), the moving average models (MA), and the autoregressive moving average models (ARMA).

3.1 Stable Time Series Forecasting Model

(1) Autoregressive models (AR)

In AR, the current observation in a series is expressed as a linear function of p previous observations, a constant term, and a disturbance term, and is expressed as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \delta + \varepsilon_t \tag{2}$$

where y_t, $y_{t-1}, \ldots, y_{t-p}$ are the observations in periods t, t–1, …, t–p, p is the number of periods (lags) considered in the development of the model, $\emptyset_i$ are the autoregressive parameters, δ is a constant term, which is related to the mean of the process; ε_t is the disturbance for period t (this model is written as AR(p)).

(2) Moving average models (MA)

In the MA of order q it is assumed that the current observation is the sum of the current and weighted past disturbances as well as a constant term

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{3}$$

where y_t is the observation in period t, q is the number of periods (order), θ_i are the moving average parameters, u is the constant term, and ε_t is the random disturbances term for periods t, t −1, …, t − q (this model is written as MA(q)).

(3) Autoregressive moving average models (ARMA)

Furthermore, there are time series models that have both autoregressive and moving average terms. These models are written as ARMA(p, q) and have the following general form:

$$\phi(B) y_t = \delta + \theta(B) \varepsilon_t \tag{4}$$

3.2 Unstable Time Series Forecasting Model

For unstable time series, although the overall level of data fluctuation does not behave in different periods of time, under the premise of different levels, the generalized features have similarity, which can be represented by a generalized autoregressive operator φB, the operator can be expressed as:

$$\phi(B) = \phi(B)(1 - B)^d \tag{5}$$

where $\phi(B)$ is an autoregressive operator. Thus, a model of homogeneous unstable time series can be expressed as follows:

$$\phi(B)Z_t = \phi(B)(1-B)^d Z_t = \theta(B)a_t \tag{6}$$

$$\phi(B)w_t = (1-B)^d Z_t \tag{7}$$

Developed and discussed throughout this chapter, the usual notation for an ARIMA model is ARIMA (p, d, q) where p refers to the models' autoregressive order, q refers to the moving average order, and d refers to the degree of differencing needed to achieve stationarity. When a series is stable, ARIMA models become ARMA. The general form of ARIMA (p, d, q) model is as follows:

$$\phi_P(B)(1-B)^d y_t = \theta_0 + \theta_q(B)a_t \tag{8}$$

3.3 Forecasting Process of Parking Berth Demand in Residential Areas

Based on the basic time series model, the process of forecasting the parking berth demand in residential area is as follows:

Step 1: Stability test on time series

Take test on the time series data of parking berth in the observed residential area. If it passed the test, then turn to the next step. Otherwise, the data is subjected to a d-order differential operation until the data is a stable time series. The David Dickey and Wayne Fuller (DF) test for unit roots is the most commonly used method in the process of testing the stability of data. The principle is as follows:

$$y_t = \rho y_{t-1} + \mu_t \tag{9}$$

where u_t is the random error of white noise (zero mean, constant variance, non-autocorrelation). From Formula (9):

$$y_{t-1} = \rho y_{t-2} + \mu_{t-1} \tag{10}$$

$$y_{t-1} = \rho y_{t-1-1} + \mu_{t-1} \tag{11}$$

Put the right side of the equation on the left, the equation can be expressed as

$$y_t = \rho^1 y_{t-1} + \rho\mu_{t-1} + \rho^2\mu_{t-2} + \cdots + \rho^1\mu_{t-1} + \mu_t \tag{12}$$

For Formula (9), the DF test is the significance test of its coefficients. The null hypothesis condition to be established is: H_0:$\rho = 1$. If the null hypothesis is rejected,

it is said that y_t has no unit root and it is stable, if not, y_t is called a random walk series, and it is unstable. Formula (9) can also be written as:

$$\Delta Y_t = (\rho - 1) y_{t-1} + \mu_t = \delta y_{t-1} + \mu_t \tag{13}$$

where $\Delta y_t = y_t - y_{t-1}$, Δy_t is a first order difference operator.

Usually Formula (13) is modified as follows:

$$\Delta Y_t = \delta Y_{t-1} - \alpha_i \sum_{i=1}^{m} \Delta Y_{t-i} + \mu_i \tag{14}$$

The DF test, which is based on Formula (14), is also called the Augumented Dickey-Fuller (ADF). If the calculated statistic value of τ is less than the value in the ADF distribution table, the series is stable.

Step 2: Calculate autocorrelation and partial autocorrelation coefficient

It is the basis to determine the model type and order to calculate autocorrelation coefficient (AC) and partial autocorrelation coefficient (PAC). The AC of stable data has a feature of censoring or trailing after the time series passing the stability test (censoring refers to a lag period after all the autocorrelation coefficient is 0, while trailing refers to the autocorrelation coefficient decreases rapidly and tends to 0 as the lag increases).

Step 3: Model identification

Select the appropriate ARMA model according to the properties of the series autocorrelation coefficient (AC) and the partial autocorrelation coefficient (PAC). When the AC tails off exponentially to 0, the model is AR and its order is determined by the number of significant lags in the PAC. When the PAC tails off exponentially to 0, the model is a MA, and its order is determined by the number of statistically significant lags in the AC. When both the AC and PAC tail off exponentially to 0, the model is ARMA.

Step 4: Parameter Estimation

The fourth step is estimating model parameters. After selecting a model based on its AC and PAC, its parameters are estimated by maximizing the corresponding likelihood function.

Step 5: Validity checking

After completing the estimation of the model parameters, the validity of model and parameters need to be tested separately. If the significance test is passed, then turn to next step; Otherwise, return to the model identification step, remodel the order and the scale of the model until the test is passed.

Step 6: Parking berth demand forecasting

When the ARIMA (p, d, q) model is used for short-term prediction, the original unstable series $y_t (y_t \sim I(d))$ is first carried through d-order difference, which is converted into the stable series W_t, that is,

$$W_t = (1 - B)^d y_t \tag{15}$$

where $W_t \sim I(0)$, then the ARMA (p, q) model is established for the stable series and the prediction of W_t is achieved as:

$$\widehat{W}_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \tag{16}$$

In general, the traffic flow data series can be transformed into a stable series by first-order difference. Therefore, take the first-order difference as an example, the prediction of original series Z_t can be described as:

$$\hat{y}_t = y_{t-1} + \widehat{W}_t \tag{17}$$

where $\hat{y}_t$ is the predicted value of original series at time interval t; y_{t-1} is the observed value of original series at time interval $t - 1$; $\widehat{W}_t$ is the predicted value of the first-order difference series at time interval t, and

$$\widehat{W}_t = \phi_1 W_{t-1} + a_t - \theta_1 a_{t-1}.$$

4 Case Study

In this paper, we selected three representative typical residential areas in the central area of Yangzhou City to do empirical analysis of the adaptability and accuracy of the model, and to verify the time-varying characteristics of parking berth demand in residential area on weekdays and weekends.

4.1 Data Sources

According to Formula (1), the total number of real-time berth occupancy of the three residential areas on weekdays was 1440, including 864 data in the first three days as the model estimation, so that 267 the predicted amount of data was ($1440 - 864 = 576$). Over the weekend, there were 1152 data collected 268 for 4 days, including

864 data in the first 3 days as model estimation, which was used to forecast data 269 volume (1152 − 864 = 288).

4.2 Model Building

According to the forecasting process of parking berth in residential area introduced above, respectively, the three residential areas' forecasting model of parking berth demand on weekdays and weekends are built as shown in Table 1.

4.3 Precision Test

Because forecasting is often the primary goal of time series analysis, it is imperative that predictive accuracy be assessed. Usually, accuracy implies how well the model reproduces the already known data. In transportation modeling, commonly used measures for evaluating the accuracy of the forecasting models are the mean square error (MSE), the mean absolute deviation (MAD), and the mean absolute percent error (MAPE).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|X_i - \widehat{X}_i\right| \tag{18}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_i - \widehat{X}_i}{X_i}\right| \tag{19}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \widehat{X}_i\right)^2}{n}} \tag{20}$$

where n is the number of samples; X_i is the actual observed value; $\widehat{X}_i$ is the predictive value.

As can be seen from Table 2:

(1) The results of the real-time forecasting model of parking berth in the three areas on weekdays and weekends have reached the high precision (100%-MAPE) more than 95%, which reflects the choice of this category. The accuracy of the six prediction models is higher than 95%, which reflects the parking berth demand in the residential areas we selected has a strong regularity of time variation.
(2) From the statistical results of the sub-period, the prediction accuracy of the above six models for the 0:00–7:59 period is higher than that of other periods due to the relatively small number of vehicles departure during the period from

Table 1 The forecast model of parking berth demand in the three areas on weekdays and weekends

Area No.	Data resource	Series type	t_δ	ADF	Stability	Model recognition	Model parameters	Prediction accuracy (%)
1	Weekdays	Raw first-order difference	−2.69484 −7.88882	−2.86470	Stable Unstable	AR1MA (6,1,0)	C = 0; Ø1 = 0.0827; Ø2 = 0.1370; Ø3 = 0; Ø4 = 0; Ø5 = 0.1304; Ø6 = 0.1270	98.68
	Weekends	Raw	−4.57290	−2.86468	Stable	ARIMA (0,0,1)	C = 81.3209; θ1 = 0.8533	97.08
2	Weekdays	Raw first-order difference	−2.75481–9.80533	−2.86469	Unstable Stable	ARIMA (6,1,0)	C = 0; Ø1 = 0; Ø2 = 0.1365; Ø3 = 0.1677; Ø4 = 0.1043; Ø5 = 0; Ø6 = 0.0813	98.41
	Weekends	Raw	−4.57290	−2.86468	Stable	ARIMA (0,0,1)	C = 69.9604; θ1 = 0.9455	96.97
3	Weekdays	Raw first-order difference	−2.10020 −8.64258 −1.82513	−2.86470	Unstable Stable	ARIMA (5,1,0)	C = 0; Ø1 = 0; Ø2 = 0.1183; Ø3 = 0.1197; Ø4 = 0.1196; Ø5 = 0.1136	98.81
	Weekends	Raw first-order difference	−29.57945	−2.86468	Unstable Stable	ARIMA (8,1,0)	C = 0; Ø1 = 0; Ø2 = 0; 03 = 0; Ø4 = 0: Ø5 = 0; Ø6 = –0.0725; Ø7 = 0.0796; Ø8 = 0.0680	98.97

Table 2 Statistical table of forecast model accuracy of parking berth demand in residential areas

Area No.	Forecast time	Forecast model	Overall performance			Sub-period		
			MAE	MAPE(%)	RMSE	RMSE of peak		RMSE of off-peak
						0:00–7:59	16:01–23:59	8:00–16:00
1	Weekdays	ARIMA (6,1,0)	0.8283	1.3159	1.2831	0.8077	1.5452	1.3734
	Weekends	ARIMA (0,0,1)	2.2674	2.9158	3.0097	1.8527	4.1326	2.5818
2	Weekdays	ARIMA (6,1,0)	0.9314	1.5932	1.4566	0.7776	1.6735	1.7019
	Weekends	ARIMA (0,0,1)	2.3122	3.0340	3.1107	1.8336	4.2114	2.2316
3	Weekdays	ARIMA (5,1,0)	0.6835	1.1890	1.0448	0.8237	1.1315	1.1418
	Weekends	ARIMA (8,1,0)	0.7740	1.0256	1.0894	0.7514	1.3331	1.1038

midnight to early morning. Relatively, the standard deviation of model in the off-peak period of parking demand is only floating between 1% and 3%. It can be concluded that the model can accurately predict the occupancy of real-time berth in residential areas from 8:00 am to 1:00 pm every day.

5 Conclusions

Based on the time-varying characteristics and the prediction method of parking berth demand in residential areas, some conclusions have been drawn.

(1) In order to explore the time-varying characteristic of parking berth demand more precisely, this paper analyzed historical data of parking berth occupancy, motor vehicles arrival, and vehicles departure respectively on weekdays and weekends which were obtained from the three residential areas.

(2) From the survey results of motor vehicles arrival and departure in residential areas, it is intuitively found that the change of parking berth in residential areas on weekdays has a more regular tendency: throughout the day, there are two peaks of vehicles arrival and three departure peaks, the vehicles arrival at the rest of time is stable. Relatively, there is no significant characteristics of vehicles arrival in residential areas on weekends, it also cannot be observed obvious peak value of vehicles departure in residential areas on weekends.

(3) Based on the time series method, considering the obvious difference between the characteristics of motor vehicles arrival and departure in residential areas on weekdays and weekends, therefore, ARIMA model and ARMA model were respectively established to forecast the berth occupancy states of three residential

areas on weekdays and weekends. The results showed that: due to the higher proportion of commute travel on weekdays and the higher proportion of elastic travel on weekends, the real-time berth occupancy on weekdays can be well forecasted by ARIMA model, which can reach more than 98% of the prediction accuracy. For weekends, the changing trend of berth occupancy is not as stable as that on weekdays, thus choosing ARMA model to forecast, and the prediction results can reach a higher accuracy of more than 95%.

In general, the time series prediction method has good adaptability to the residential area, and the higher accuracy can be achieved by selecting the appropriate model. However, there is a large margin of improvement for future research. This research illustrated the time-varying characteristics of parking berth demand only in a certain type of residential areas but did not study on any other kinds of residential areas, which is to be explored in the next stage.

Acknowledgements The research is supported by the Project of National Natural Science Foundation of China (Grant No. 51478111); Shanghai Rising-Star Program (16QB1403000); Shanghai Urban-Rural Development Transportation Talents Special Funds.

References

1. Levinson HS (1982) Parking in a changing time. Publication of Newcastle University Australia
2. Vlahogianni EI, Kepaptsoglou K, Tsetsos V et al (2015) A real-time parking prediction system for smart cities. J Intell Transp Syst 20(2)
3. Chen K, Wang JJ, Han F (2012) Research of parking demand forecast model based on regional development. In: The twelfth COTA international conference of transportation professionals, 23–29
4. Xue XJ, Ou XQ, Yan KF (2010) Parking demand forecasting for space sharing facility in new urban area. Urban Transp China 8(5):53–55
5. Qin HM, Guan HZ, Sun WL, Xiao Q (2011) Study of the parking shared demand model of urban mixed use lands—Hua Mao center in Beijing as an example. J Beijing Univ Technol 08:1184–1189
6. Guan HZ, Wang X, Wang X (2006) The research on forecasting method for parking demanding. J Beijing Univ Technol 32(7):600–604

Risk Evaluation Model of Unsignalized Intersection Based on Traffic Conflict Line Theory

Li Yuan, Yi-hang Sun, Xuan Zhang and Juan He

Abstract In order to analyze unsignalized intersections safety performance objectively and systematically, given that the traditional traffic conflict research focuses on the two-vehicle conflict, a risk evaluation model based on "traffic flow conflict line" theory was developed. Firstly the characteristics of crossing, diverging and merging conflicts were analyzed. Due to different types of traffic conflicts caused different potential treats and collision severities, the research estimated the weighted values of crossing, merging and diverging conflicts through analyzing the leading vehicle conflict potential probability, collision severities and transmission length of traffic conflicts. Finally the research combined all results for establishing useful index of intersection safety called expected values of conflicts. Compared with the traditional method, the results show that the leading vehicle conflict probability model is closer to the real process of traffic conflicts based on the critical conflict distance. The model takes into account the speed, angle, acceleration and reaction time between two vehicles. Based on physics collision theory, analyzing the angle change, deceleration relation of the vehicle and the kinetic energy loss, three traffic conflicts weights were discussed when accident occurred (crossing conflict: diverging conflict: merging conflict = 12.7051:1:1). Based on the mathematical expectation method, the total equivalent expected conflict model was developed through analyzing the potential probability of conflict, traffic volume and vehicle position, so this model can more accurately describe the actual traffic flow conflict behavior.

Keywords Traffic engineering · Traffic flow conflict line · Unsignalized intersection · Traffic safety

L. Yuan · Y. Sun (✉) · J. He
College of Civil and Transportation Engineering, Hohai University, Nanjing Jiangsu 210098, China
e-mail: pi9900@126.com

L. Yuan
e-mail: yuanlibox@hhu.edu.cn

X. Zhang
School of City and Architecture Engineering, Zaozhuang University, Zaozhuang Shandong 277160, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_5

1 Introduction

Different traffic flow interweaves at the intersection and it is easy to lead to the occurrence of traffic accidents. According to the statistics, approximately 47% of the total 10,064,000 crashes in the United States in 2013 occurred at intersections. The American Institute of Transportation Engineers have declared that road intersection safety was an important subject that required careful solutions [5]. At unsignalized intersection, the probability of the occurrence of traffic accidents is higher. One reason is the lack of unified coordination and systematic traffic control. The other reason is that the traffic behavior is only constrained and guided by traffic signs and markings. Furthermore, part of drivers are not in accordance with the traffic rules while they are driving.

By evaluating the risk of unsignalized intersection, traffic accidents can be effectively prevented. The number of traffic accidents can be reduced and the severity of traffic accidents can be controlled. Based on the traffic conflict technology, many evaluation systems have been proposed. Poul Greibe proposed accident forecasting models for three-way intersection, four-way intersection, signalized intersection and unsignalized intersection [3]. Dagmar presented the project KONFLIKT which aims develop a methodology of observation and evaluation of traffic conflicts and its application of traffic conflict techniques in Czech Republic. By using two-stage model in British Columbia, Karim found that there was a relationship between traffic conflict and accident. He proposed that the probability of the occurrence of traffic conflict in the urban intersection was much higher than that in the suburban intersection. Zhou presented a new way to pre-evaluate the safety performance of intersection improvement strategies using VISSIM and SSAM, and the procedure was exemplified by a signalized intersection. Shahdah has presented a model for integrating observed crash-based and simulated conflict-based indicators to obtain treatment Crash modification factors (CMF) [7]. Dorinela proposed that traffic conflict index can be used as a useful measure to compare different configuration and systematization schemes [2]. Yin and Zou defined conflict risk degree as the index of evaluating the impacts of traffic accidents on intersection safety, and a calculation model involving conflict probability, influencing degree and encountering time was developed as well [9]. Huang and Liu developed conflict prediction models using the simulated conflicts as independent variables. The calibrated simulation models provided reasonable estimates for rear-end and total conflicts [4]. Wael K. M. Alhajyaseen developed the conflict index for safety assessment of intersections considering crash probability and severity [1].

To some degree, existing evaluation methods can meet the basic requirement of effectiveness. But the evaluation index is relatively single because it only considers the conflict between two vehicles. It can not roundly reflect the safety status of the intersection. Based on the national natural science fund project ‘The formation mechanism of intersection traffic flow line and the research on safety risk evaluation model’, this paper studies the influence on following vehicles that caused by the conflict between the two vehicles which are in front of them in the intersection.

By using multiple indexes, the proposed risk evaluation model in this paper can be used to evaluate the safety performance of unsignalized intersection comprehensively. According to the proposed model, the safety level and management level of unsignalized intersection can be improved.

2 Traffic Flow Conflict Line Theory

Existing traffic conflict technology is only based on the behavior that happened between two vehicles. But in general, it is impossible that there are only two vehicles on the road. So within a certain time or distance, the front conflict will affect the following vehicles to some extent. If the following drivers could not make correct choice, it would easily lead to the occurrence of rear-end accidents.

As shown in Fig. 1, there is a conflict between vehicle B and vehicle A at the intersection. In order to avoid the occurrence of collision, vehicle A will take some measures, such as braking, deceleration and so on [8]. If vehicle C followed vehicle A, then within a certain distance, the behavior of vehicle A would affect vehicle C. If vehicle C didn't take some measures, it might lead to a collision. So vehicle C must take some measures. This status will pass on through the following vehicles. The line which was composed of the affected following vehicles was called as 'traffic flow conflict line'.

As shown in the definition, the formation of traffic flow conflict line has two premises. First of all, there is a conflict between two vehicles at the intersection. Second, the conflict leads to the occurrence of vehicle queuing. Both of them are essential and they interact with each other. Conflict leads to vehicle queuing and

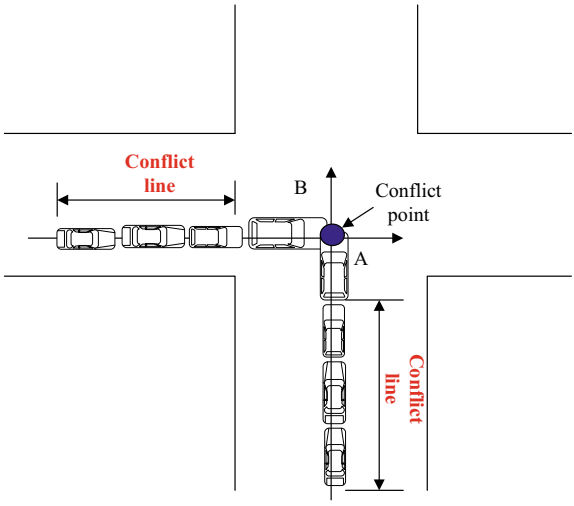


Fig. 1 Traffic flow conflict line

vehicle queuing will lead to the occurrence of new conflict. Until the conflicts evanish, the traffic system returns to normal.

The larger the extent of traffic flow conflict line is, the larger the number of queuing vehicles will be. It will not only cause congestion, but also cause great traffic risk, sometimes even lead to crash. But when the extent exceeds a threshold value, in consideration of safety, drivers will decelerate. At this time, there is a negative correlation between the number of traffic conflict and traffic volume. So the quantity of traffic volume has a great effect on the number of traffic conflicts.

On the basis of the analysis of the whole process of traffic conflict, certain mathematical methods can be used to get the reach of traffic flow conflict line. The longer the traffic flow conflict line is, the more the number of affected vehicles is. It means that the severity of the conflict is high and even leads to severe rear-end accidents.

3 Develop Risk Evaluation Model

As shown in Fig. 2, on the basis of the analysis of traffic conflict, the risk evaluation model was developed.

$$Y = f(P_i, X_i, N_i) \tag{1}$$

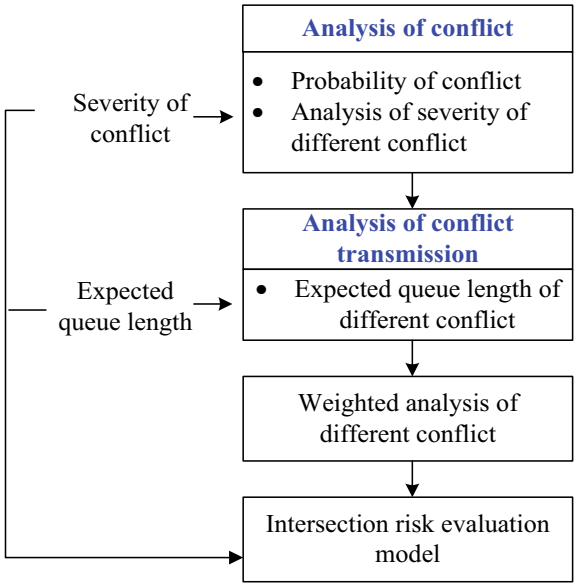


Fig. 2 The establishing process of intersection risk evaluation model

where, Y is the total expected values of conflicts, P_i is the probability of different traffic conflicts, X_i is the collision severities of different traffic conflicts, N_i is the expected values of different conflicts.

3.1 Analysis of Collision Severity Ratio

Different types of traffic conflicts cause different potential crash rate and different collision severities. Based on the probability of conflicts and collision severities, the collision severity ratio was proposed.

$$X_i = P_i B_i (i = 1, 2, 3) \tag{2}$$

where, X_1, X_2, X_3 are the collision severity ratio of crossing, merging, and diverging conflicts respectively; P_1, P_2, P_3 are the probability of crossing, merging, and diverging conflicts respectively; B_1, B_2, B_3 are the collision severities of crossing, merging, and diverging conflicts respectively.

3.1.1 Traffic Conflict Probability Model

In the process of driving, there is a critical distance between two vehicles. Once the two vehicles' distance is less than the threshold, the drivers will become stressed and their psychology and behavior will be influenced. At this time, the drivers must take some measures to avoid the occurrence of collision such as braking, turning and so on. The threshold is called as critical conflict distance.

As shown in Fig. 3, considering vehicle as particle, a rectangular coordinate system was established to analyze the traffic conflict. The direction of the vehicle I is used as y axis forward direction. As shown in Fig. 2, V_1 is the speed of vehicle I, V_2 is the speed of vehicle II, V_{21} is the relative speed, ψ is the angle between V_{21} and x axis forward direction, t_1 is the time from pedaling to the generation of braking

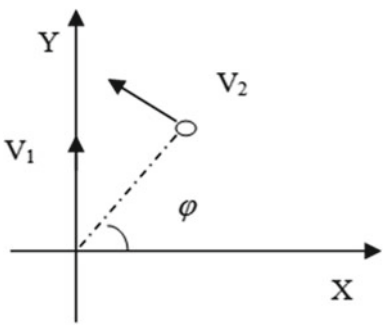


Fig. 3 Analysis of the critical particle conflict

force, t_2 is the time of the process of braking force increasing, a_2 is the acceleration of vehicle II, t_0 is the drivers' reaction time. This paper gets the formula of critical conflict distance by using the formula of vehicle braking distance and meanwhile taking t_0 into consideration, where $t_0 = t_0' + t_0'' + t_0'''$.

$$R(\phi) = V_2 \frac{\cos\left|\phi - \arcsin\left(\left|\frac{V_1}{V_2}\right| \cos\phi\right)\right|}{\cos\phi} \left(t_1 + \frac{1}{2}t_2 + \frac{|V_2|}{2|a_2|} + t_0 \cos\phi\right) \tag{3}$$

where, t_0' is the time that drivers spend to be aware of danger and tend to take measures, t_0'' is the time that drivers spend to move their feet to the brake pedal, t_0''' is drivers' reaction-error time.

When two vehicles close to each other, there is a critical conflict distance between them. So drivers need to use their observation and experience to judge their space with other vehicles and determine if they should implement a certain behavior. This process can be considered as gap acceptance behavior. Logit model [6] can be used to describe this process:

$$\begin{cases} P_a = \dfrac{\exp\left(-\beta_0 + \sum_{i=1}^{k} \beta_i x_i\right)}{1 + \exp\left(-\beta_0 + \sum_{i=1}^{k} \beta_i x_i\right)} \\ P_r = 1 - P_a \end{cases} \tag{4}$$

where, P_a is the probability of acceptance, P_r is the probability of rejection, x_i is influence factor, β_i is parameter.

Through field investigation and analysis, the relative speed and spacing are found to be the main factors that affect drivers' behavior. In equation set (4), β_i are unknown parameters. The method of maximum likelihood estimation can be used to calibrate them.

Define y_i:

$$y_i \begin{cases} 1 \ldots \textit{Behavior within critical conflict distance for the i time } (R < R(\phi)) \\ 0 \ldots \textit{Behavior over critical conflict distance for the i time } (R > R(\phi)) \end{cases} \tag{5}$$

For the i time conflict, the speed of vehicle A and vehicle B and their space should be recorded. The following likelihood function is obtained by (4):

$$L^* = \prod_{i=1}^{N} P_a^{y_i} P_r^{1-y_i} \tag{6}$$

The value of β_i can be calculated by Newton-Raphson. According to β_i, P_a can be calculated. P_a is leading vehicle's probability of traffic conflict.

The greater the probability of traffic conflict is, the greater the probability of traffic accident might be. But the great probability of traffic conflict doesn't mean that the severity of conflict is great. So the severity of conflict should be analyzed.

3.1.2 Analysis of Collision Severity

There are three kinds of traffic conflicts at the intersection: crossing conflict, merging conflict and diverging conflict. Different type can lead to different collision severity. This section introduced the collision theory in physics. By calculating the loss of kinetic energy caused by collision, the different severity weights of different conflicts were obtained.

As shown in Fig. 4, M_A is the quality of vehicle A, V_A is the speed of vehicle A, M_B is the quality of vehicle B, V_B is the speed of vehicle B, θ_1 is the angle between vehicle A and x axis, θ_2 is the angle between vehicle B and x axis. After collision, vehicle A and vehicle B turn into a whole. Its quality is $M_A + M_B$. V_C is its speed and α is the angle between it and x axis. When the collision occurs, the loss of kinetic energy can be calculated by the following formula ($0 \le \theta_1 \le \pi/2$, $0 \le \theta_2 \le \pi/2$):

$$\Delta W = \frac{\int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \frac{MV^2}{2}(1 - \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2) d\theta_1 d\theta_2}{\int_0^{\frac{\pi}{2}} 1 d\theta_1 \int_0^{\frac{\pi}{2}} 1 d\theta_2} \tag{7}$$

Convert it into:

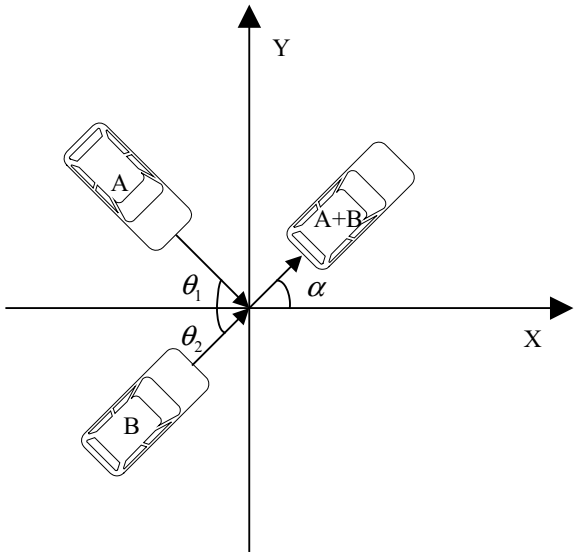


Fig. 4 Cross conflict

Table 1 Weight value of expected traffic conflict calculation

Type	B_i	P_i	$X_i = B_i P_i$	$X_i = \frac{3X_i}{\sum X_i}$
Crossing	12.7051	P_1	$12.7051P_1$	$\frac{38.1183P_1}{\sum X_i}$
Merge	1	P_2	P_2	$\frac{3P_2}{\sum X_i}$
Diverge	1	P_3	P_3	$\frac{3P_3}{\sum X_i}$

$$\Delta W = \left(\frac{1}{2} - \frac{4}{\pi^2}\right) MV^2 \tag{8}$$

For merging conflict and diverging conflict, the same method can be used to calculate their loss of kinetic energy ($0 \leq \theta_1 \leq 45°$, $0 \leq \theta_2 \leq 45°$).

$$\Delta \mathrm{W} = \frac{\int_0^{\frac{\pi}{4}} \int_0^{\frac{\pi}{4}} \frac{\mathrm{MV}^2}{2} (1 - \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2) d\theta_1 d\theta_2}{\int_0^{\frac{\pi}{2}} 1 d\theta_1 \int_0^{\frac{\pi}{2}} 1 d\theta_2} \tag{9}$$

Convert it into:

$$\Delta W = \left(\frac{1}{2} - \frac{16}{\pi^2} + \frac{8\sqrt{2}}{\pi^2}\right) MV^2 \tag{10}$$

According to the above calculation process, the different severity weights can be obtained. The calculation result was 12.7051:1:1 (B_1:B_2:B_3).

3.1.3 Establish Collision Severity Ratio Model

As shown in Table 1, in order to avoid distorting the calculation, the sum weight of the three types of conflicts was three.

3.2 Analysis of Transmission Length Model of Traffic Conflict

The transmission length of traffic conflict can be defined as the total number of vehicles influenced by conflicts. Based on the mathematical expected value theory, the comprehensive analysis of the probability of conflict, traffic volume and vehicle location, the different transmission length models were established respectively.

Because the difference between crossing conflict and merging conflict was only the angle of conflict, the construction of the models were discussed together.

When there was crossing or merging behavior between X vehicles and N vehicles, the number of position which could be used by the X vehicles was $N + 1$. On the

basis of permutation and combination, the random permutation number of the $N + X$ vehicles was *(N + X)!* Because the behind vehicles wouldn't cross or merge with the front vehicles, so the number of the practical permutation was L and the probability was P.

$$L = \frac{(N+X)!}{N!X!} = C_N^{N+X} \tag{11}$$

$$P = \frac{1}{L} = \frac{1}{C_N^{N+X}} \tag{12}$$

$$TC_x = \sum_{q_x=0}^{N} \sum_{q_{x-1}=0}^{q_x} \cdots \sum_{q_1=0}^{q_2} (q_1 + q_2 + \cdots + q_x) = \frac{X(X+1)}{2} \cdot C_{X+1}^{N+X} \tag{13}$$

Assumed that the probability of each situation was equal, the expected value of crossing or merging conflict could be calculated.

$$\begin{aligned} E(TC_x) = TC_x \times P &= \frac{X(X+1)}{2} \cdot C_{X+1}^{N+X} \cdot \frac{1}{C_X^{N+X}} \\ &= \frac{X(X+1)}{2} \cdot \frac{(X+N)!}{(X+1)!(N-1)!} \cdot \frac{X!N!}{(X+N)!} = \frac{NX}{2} \end{aligned} \tag{14}$$

In the same way, based on the expected value theory ($E(aX) = aE(X)$), the expected value of diverging conflict could be calculated.

$$E(X \cdot TC_x) = X \cdot E(TC_x) = X \cdot \frac{N-1}{2} = \frac{(N-1)X}{2} \tag{15}$$

3.3 Develop Risk Evaluation Model

Based on the collision severity ratio and transmission length of conflict, the risk evaluation model was developed.

$$Y = \sum_{i=1}^{3} N_i X_i = \frac{38.1183 P_1}{\sum X_i} N_1 + \frac{3P_2}{\sum X_i} N_2 + \frac{3P_3}{\sum X_i} N_3 \tag{16}$$

where, Y is the total expected values of conflicts.

Table 2 Risk level of intersection

Risk level	*Y*
Safe	≤600
Generally safe	≤1000
Generally dangerous	≤1450
Dangerous	>1450

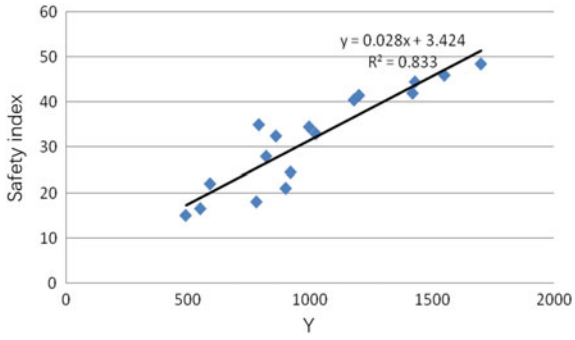


Fig. 5 Comparison between model rating and safety index rating

3.4 Classification of Risk Level and Model Validation

Y was used to be the index of the risk classification.

In order to divide the intersection risk level, traffic data was collected based on field investigation in Shandong Province and Jiangsu Province. According to the intersection safety situation and combined with the researchers' judgement, as shown in Table 2, the risk classification of intersection was divided into four levels.

In order to verify the risk evaluation model, the method of security index was adopted. The method was using the hidden danger significance and the potential accident severity to establish a safety diagnostic ranking model.

Traffic data was collected in Shandong Province and Jiangsu Province. Then the security index was calculated. As shown in Fig. 5, MATLAB was used to make linear regression of the security index and *Y*. The correlation coefficient was 0.83.

4 Case Study

As shown in Fig. 6, the unsignalized intersection which is between Shanguo South Road and Shanwen Street was chosen as an example in this paper. The south-north approach of the intersection is Shanguo South Road and it is an urban two-way four-lane side road. The east-west approach is Shanwen Street and it is an urban two-way two-lane branch road. Each lane's width is 3.25 m. There is no channelization at the intersection.

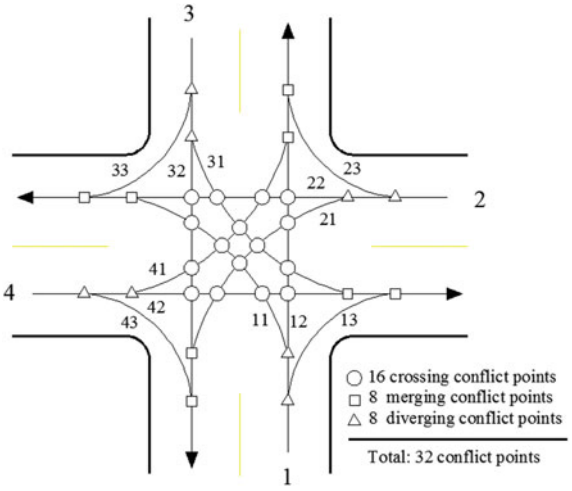


Fig. 6 Traffic conflict point and type for intersection

Step 1: Data acquisition.

The traffic volume is shown in Table 3.

During the survey period, the number of traffic conflicts was 264, and 259 of them were recorded. As shown in Table 4, the number of conflicts that occurred during peak hours was much bigger than that of off-peak hours.

The data of the peak hour (17:30–18:30) was chosen to evaluate the intersection's risk level.

Step 2: Calculation of conflict probability of the leading vehicle.

The speed of vehicles was obtained by radar. According to Eq. (3), critical conflict distances of different speed were calculated. Then based on the gap acceptance theory, if the two vehicles within the critical conflict distance or not can be calculated and judged. Bringing the results into Eq. (6) can obtain the likelihood function. Solving the maximum likelihood function can obtain the three kinds of conflict probability. The results are shown in Table 5.

Step 3: Calculation of transmission length (Table 6).
Step 4: Calculation of expected value of conflicts.

$$Y = \sum_{i=1}^{3} N_i X_i = 2.538 \times 1923 + 0.197 \times 717 + 0.265 \times 1429 = 5400(\text{pcu/min})$$

According to Table 2, the risk level of the intersection is dangerous. It means that the intersection is unsafe. There is no channelization at the intersection. In order to improve the safety level of the intersection, channelization should be advised to set up.

Table 3 Average traffic volume per minute of intersection

South			East			North			West		
Left (N_{11})	Straight (N_{12})	Right (N_{13})	Left (N_{21})	Straight (N_{22})	Right (N_{23})	Left (N_{31})	Straight (N_{32})	Right (N_{33})	Left (N_{41})	Straight (N_{42})	Right (N_{43})
5	9	4	4	21	5	6	18	6	7	18	4

Table 4 The number of traffic conflicts

Time	Number	Merging	Diverging	Crossing
8:00–9:00	98	21	45	32
9:30–10:30	20	5	6	9
15:00–16:00	28	6	12	10
17:30–18:30	113	22	52	39

Table 5 Conflict probability

Conflict type	Crossing conflict	Merging conflict	Diverging conflict
Probability	0.491	0.483	0.652

Table 6 Total expected conflict amount for each approach

Type	North	West	South	East	Total
Crossing	582	499	346	496	$N_1 = 1923$
Merging	216	183	170	148	$N_2 = 717$
Diverging	435	406	153	435	$N_3 = 1429$

5 Conclusion

Based on the traffic flow conflict line theory, a risk evaluation model of unsignalized intersection was proposed in this paper. Based on the conflict probability of the leading vehicle, severity of conflict and expected transmission length, the risk level of unsignalized intersections was evaluated. There are some conclusions:

(1) According to the reality at the intersection, the traffic flow conflict line theory was proposed and its generation mechanism and characteristics were analyzed;
(2) Considering probability, severity and expected transmission length of different conflicts, a risk evaluation model of unsignalized intersection based on traffic conflict line theory was proposed. The index of intersection safety which was called expected values of conflicts was used to describe unsignalized intersections' risk level.

There is also something that needs to be researched and improved further:

(1) This paper mainly analyzed problems from the view of people and vehicles, lacking the view of road and environment;
(2) This paper only studied unsignalized cross intersections. It remains to be discussed if the risk evaluation model could be applied to other types of unsignalized intersections.

Acknowledgements This research was supported in part by the National Natural Science Foundation of China with no. 51308192, and by Science Foundation of Ministry of Education of China with no. 12YJCZH062. The authors would like to show great appreciation for these supports.

References

1. Alhajyaseen WKM (2014) The development of conflict index for the safety assessment of intersections considering crash probability and severity. Procedia Comput Sci 32:364–371
2. Costescu D, Raicu S, Rosca M et al (2016) Using intersection conflict index in urban traffic risk evaluation. Procedia Technol 22:319–326
3. Greibe P (2003) Accident prediction models for urban roads. Accid Anal Prev 35(2):273–285
4. Huang F, Liu P, Yu H et al (2013) Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections. Accid Anal Prev 50:1014–1024
5. Liu M, Chen Y, Lu G et al (2017) Modeling crossing behavior of drivers at unsignalized intersections with consideration of risk perception. Transp Res Part F Traffic Psychol Behav 45:14–26
6. Savolainen PT (2016) Examining driver behavior at the onset of yellow in a traffic simulator environment: comparisons between random parameters and latent class logit models. Accid Anal Prev 96:300–307
7. Shahdah U, Saccomanno F, Persaud B (2014) Integrated traffic conflict model for estimating crash modification factors. Accid Anal Prev 71:228–235
8. Treiber M, Kesting A, Helbing D (2006) Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps. Rev Part E, Phys, p 74
9. Yin W, Zou Q, Lv C et al (2016) A quantitative model of conflict risk degree at non-signalized intersections. Procedia Eng 137:171–179

Understanding the Impacts of Leisure Purpose and Environmental Factors on the Elders Leisure Activities and Travel Behavior: A Case Study in Kunming, China

Ren Dong, Shengyi Gao and Baohong He

Abstract In this study, authors use a structural equation modeling approach to test the impacts of leisure purpose and environmental factors on the elders' leisure features, location choice, and related travel behavior. The results showed that leisure purpose had significant direct impacts on leisure features and locations but had no direct effects on leisure related travel behavior. Environmental factors had only direct effects on leisure locations and indirect effects on leisure related travel behavior. The results suggest that community parks and neighborhood green fields were more important to the elders' leisure activities than large city parks due to high proximity and thus should be deliberately planned to improve the elders' well-being in urban land use development.

Keywords The elders · Leisure motivation · Travel behavior · Structural equation model · "Push-Pull" theory

1 Introduction

According to the World Health Organization's (WHO) report [1], everyone in the world is expected to live more than 60 years of age, and the proportion of the people aged over 60 grows faster than any other age group. This is true not only in the developed counties, but also in some developing countries. According to the China Population Statistics Yearbook [2], the population 60 years old or over has been growing significantly, accounting for 16.1% of the total population by the end of 2015. The ratios are even higher in mega cities, in this regard, China became an aging country as early as at the turn of the 21th century [3].

R. Dong · B. He (✉)
Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming 650500, China
e-mail: 94002267@qq.com

S. Gao
Sacramento Area of Council of Governments, Sacramento 94203, CA, USA

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_6

Under present national retirement policy, men are required to retire at 60 (65 for those with a senior title) and women at 55 (60 for those with a senior title). Given the high labor force participation and employment rates, almost all persons 60 years old or over in cities are retirees [4]. In China, the majority of urban retirees have pension, social security, health insurance, and potentially their own savings. In addition, they live in their own condominiums without any mortgages and do not pay real estate property taxes. They have decent and stable income to pursue a different life style from that before they retired. Obviously, the impacts of the elders, which is defined in this context as aged over 60 and not obliged to work, on society are versatile. From the view of public policies in urban planning and transportation, especially public transit, it is of interest to understand how the elders use their leisure time for activities outside their homes, where they go, and what factors affect their choices of activity locations; these factors are thus useful for policy recommendations to improve the elder's well-being.

2 Literature Review

Studies show that the elders in cities had great enthusiasm for outdoor activities and generated significantly more trips [5] and the outdoor activities increased their perception of happiness [6]. The ubiquitous collective dancing, which is called square dancing in China, is the best example of the elders' outdoor activities. By analyzing the activity diaries of 47 participants in Beijing of 122 days spanning a year, Sun et al. [5] found that the elders spent two hours on outdoor activities and the average travel distance from home was 4.7 km. The venues and neighborhood/city parks near the residences were the major activity locations. Using the third household travel survey data which was a 1% sample (i.e. 81,760 households and 208,290 persons) of the Beijing metropolitan area, Zhang [4] found that elders had more non-work trips than those of working age and the mode share of walk, bike, and transit was 58.3%, 18.3%, and 14.1%, respectively, Wang [3] surveyed 941 elders people in twelve city parks in Beijing. In this study, the elders spent 2.6 h on outdoor activities and visited parks 18.2 times per month. Sixty-three percent chose to walk to parks while 32% chose bus or subway. One-way travel time and distance averaged 35 min and 3.7 km, respectively. Guan [7] found that in Guangzhou, the most popular outdoor activity among the elders was hiking and the most popular place for leisure activity was parks. The average travel time to parks was 18 min by biking and 26 min by public transit. Wang [3] found there was a severe spatial mismatch between the spatial distributions of the elders population densities and the city parks, and the visits to the parks were affected by proximities by walk or bus.

The studies cited above provided useful information to understand the elders' leisure activities but were descriptive and did not provide statistical insight about how personal characteristics, motivations and constraints affect leisure activities, and related behavior. Crawford and Godbey [8], Crawford et al. [9], and Godbey et al. [10] proposed a hierarchical constraints model to address the relationships

between motivations for an activity and realized activities. The basic assumptions of the model were (1) that people could not participate in an activity due to constraints, which were classified as intrapersonal, interpersonal, and structural constraints, (2) that all the constraints could be explained by people's social-demographic characteristics. Based on Crawford and Godbey's model (1987), Wang [11] grouped thirty-two leisure motivation factors into four types (learning, social, physical excise, and relax) and twenty-nine leisure constraints into three types (intrapersonal, interpersonal, structural) and found that age (three age groups: 60–70, 70–80, >80) significantly affected all four leisure motivations, and income significantly affected all three leisure constraints while other variables like gender, marriage status, education, occupation, etc. did not have significant effects on the constraints at the confidence level of 0.05.

Crawford and Godbey's model (1987) did not take into account the impacts of activity-related features such as activity spatial locations, location size, aesthetics, facilities, etc. on the participation of activities. Surveys [12] showed that the activity features were very important factors to be considered in people's decision. Therefore, the hierarchical constraints model could only partially explain people's decision in activity participation. Rong [12] improved the hierarchical constraints model by incorporating nine activity-related factors which represented functionality, service level, and accessibility of activities and activity locations, and found that the smaller the constraints were, the higher the odds for the elders to participate in leisure activities would be. In other words, leisure activities which served multiple purposes and locations with good service and access would attract more elders' folks to go.

The improved hierarchical constraints model [12] shed some light on understanding how some activity-related features affected the elders' participation in leisure activities. But, it only depicted a partial picture of the elders' leisure activities. The structure of the model limited its ability to further address the relationships between the features of activity and activity location, and related travel behavior. For this reason, the authors in this paper propose a structural equation model (SEM), which is usually considered to be the best approach to understanding the complex causal relationships in behavioral studies [13], to quantify the direct and indirect effects and differentiate the roles of internal motivations and external incentives on leisure activities and related travel behavior.

3 Methodology

3.1 Data Collection

A questionnaire was designed to collect the data of the elders' activities. Besides socio-demographics, the questionnaire included six questions about leisure purpose, six questions about leisure environment cognition, three questions about the characteristics of leisure activities, and four questions about travel behaviors. The features of leisure locations were valued by the surveyors through GIS analysis. The values of

the questions were measured with a Likert five-point scale (see Table 1). The leisure purpose was identified by the major purpose of the leisure trip and a leisure trip had only one leisure purpose.

The survey was done from Sept. 11 to Oct. 29, 2017 in Kunming, Yunnan Province, China. The survey locations included neighborhood and city parks, community squares, and neighborhood green fields that spatially well represent the major activity locations of the elders. The surveys were randomly selected. The questionnaire was filled out by the investigator and the investigation together on site through a face-to-face interview. The total valid samples collected were 663. The descriptive statistics of the samples were listed in Table 2.

3.2 SEM Model

A SEM is an a priori model [14]. In the context of this study, we propose a hypothesized SEM model (see Fig. 1) to address how leisure purposes and environmental factors affect leisure activities and choices of activity location and further affect the elder's travel behavior.

In the conceptual model, six constructs (in oval circle in Fig. 1) were adopted to represent six types of variables: leisure purposes, personal attributes, environmental factors, leisure activity features, leisure locations, and travel behavior. According to literature [15–20], leisure purpose (represented by the construct *Leisure purpose* in Fig. 1) is consisted of six leisure purposes and is predicted to have direct impacts on leisure activity features (represented by the construct *Activity feature* in Fig. 1), leisure activity locations (represented by the construct *Leisure locations* in Fig. 1), and travel behavior (represented by the construct *Travel behavior* in Fig. 1). The construct *Environment cognition*, represents environmental factors and are expected to have direct impacts to leisure activities, locations, and travel behavior as leisure purpose. *Activity feature* consists of three variables *Activity content*, *Companion*, and *Activity duration*. The construct *Leisure location* represents the features of the locations of leisure activities. In the model, *Activity feature* and *Leisure location* are two endogenous variables and are hypothesized to affect directly each other. In fact, we have not found any empirical study on the interaction between activity features and locations in the literature. Besides *Activity feature* and *Leisure location*, Leisure purpose and Environmental cognition are assumed to have direct influence on travel behavior, such as travel mode, distance, schedules, and frequency. In addition, *Activity feature* and *Leisure location* have direct impacts on travel behavior (i.e. the variable *Travel behavior*). Personal attributes are assumed to have direct impacts on activity features, locations, and related travel behavior.

Table 1 Model variables

Exogenous variable

Category	Variable names	Variable symbol and value	Category	Variable names	Variable symbol and value
Leisure purpose	Physical exercise	**X1** (1–5)	Environmental cognition	Beautiful environment	**X7** (1–5)
	Visiting friends	**X2** (1–5)		Cleanliness	**X8** (1–5)
	Relaxation	**X3** (1–5)		Fresh air	**X9** (1–5)
	Hobbies	**X4** (1–5)		Perfect facilities	**X10** (1–5)
	Entertainment	**X5** (1–5)		Proximity	**X11** (1–5)
	Killing time	**X6** (1–5)		Easy access	**X12** (1–5)

Category	Variable names	Variable symbol and value
Personal attributes	Gender	**X14** (Male = 2, Female = 1)
	Age	**X15** (<66 = 1, 66–70 = 2, 71–75 = 3, 76–80 = 4, >80 = 5)
	Family status	**X16** (4 generations = 5, 3 generations = 4, 2 generations = 3, The couples = 2, Else = 1)
	Education	**X17** (Bachelor degree and above = 5, College = 4, High school and technical school = 3, Junior high school = 2, Primary school and the following = 1)
	Personal monthly income (¥)	**X18** (<1000 = 1, 1000–2000 = 2, 2000–3000 = 3, 3000–4000 = 4, >4000 = 5)

Endogenous variable

Category	Variable names	Variable symbol and value
Leisure location	Leisure site class	**Y1** (Street green space = 1, Community level = 2, Administrative district level = 3, city level = 4)
	Leisure site size	**Y2** (<5 ha = 1, 5–15 ha = 2, 20–30 ha = 3, 50–60 ha = 4, >60 ha = 5)
	Comprehensive environment	**Y3** (The level of greening rate and the level of water coverage rate, value of 2–8)
Activity feature	Activity content	**Y4** (No specific activity = 1, An activity = 2, Two activity and above = 3)
	Companion	**Y5** (Alone = 1, Accompany = 2)
	Activity duration	**Y6** (Continuous values)
Travel behavior	Travel mode	**Y7** (Walk = 1, Bike = 2, Bus = 3)

(continued)

Table 1 (continued)

Endogenous variable		
Category	Variable names	Variable symbol and value
	Departure time	**Y8** (6:30–7:29 = 1, 7:30–8:29 = 2, 8:30–12:00 = 3, 13:00–16:30 = 4)
	Travel frequency	**Y9** (Continuous values)
	Travel distance	**Y10** (<0.5 km = 1, 0.5–1 km = 2, 1–2 km = 3, 2–4 km = 4, 4–6 km = 5, >6 km = 6)

Table 2 Descriptive statistics of the sample

Attribute	Range	Percentage	Attribute	Range	Percentage
Gender	Male	48	Industry before retirement	Government	15
				Non-government public	23
	Female	52		Industrial/commercial	54
				Other	8
Age group	60–64	25	Personal monthly income (¥)	<1000	12
	65–69	32		1000–2000	32
	70–74	22		2000–3000	32
	75–79	11		3000–4000	14
	≥80	10		>4000	10
Education background	Bachelor degree and above	10	Living state	4 generations together	2
	College	13		3 generations together	23
	High school and technical school	27		2 generations together	28
	Junior high school	31		Couple	42
	Primary school or lower	19		Live alone	6

4 Results

The model parameters were estimated with the statistical software AMOS 20.0. The hypothesized model was formatted into Fig. 2 based on the significance test ($p < 0.05$) of the parameter estimates. The variables and paths in the hypothesized model

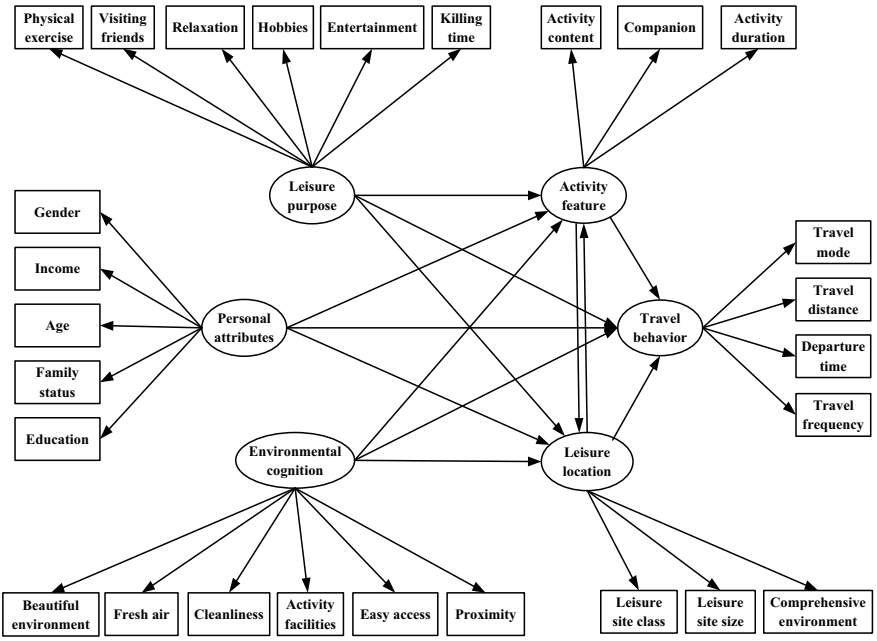


Fig. 1 Hypothesized model of the elder's leisure activity

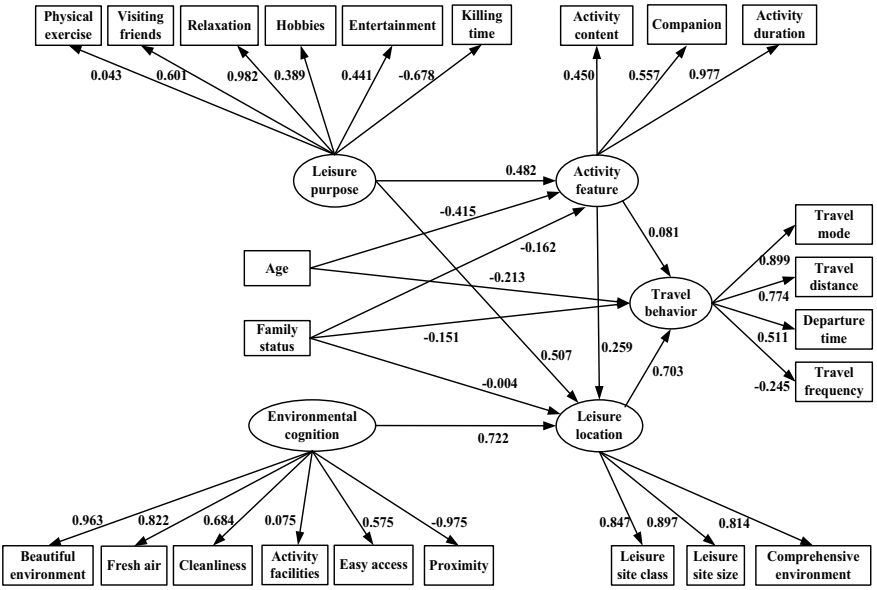


Fig. 2 Standardized parameter estimates ($p < 0.05$)

Table 3 Goodness of fit of the model

Fitting index	RMR	GFI	AGFI	PGFI
Model values	0.194	0.946	0.936	0.792
Reference values	<0.5 (Best value) <0.8 (Suitable value)	>0.9	>0.9	>0.5

were removed from the model if the *p* value of the parameter estimates was greater than 0.05, which implied the parameter estimates were statistically zero. Four widely accepted indices to measure the goodness-of-fit of SEM model were listed in Table 3. The indices suggested a good fit of the hypothesized model.

From the figures, we could see that leisure purposes, which were internal motivation and the force for people to do leisure activities, had significant positive direct effects on activity features. In other words, the leisure purposes were significant determinants in a person's leisure activity decisions about how many activities to involve, whether to have a companion, and how long to stay. *Relaxation* was the No. 1 leisure purpose and it had more impacts on activity features than other leisure purposes. *Killing time* which could be interpreted as having no clear purpose for leisure had a negative sign, which implied a negative impact on activity features and destination choices. The elders whose leisure purpose was *Killing time* tended to leisure without a companion, involve fewer activities, and spend less time outside on a trip. They were more likely to use nearby neighborhood green fields and parks other than city parks, which were usually far from residential areas. They are less picky at leisure locations than other leisure activities. As to *Physical exercise*, its impacts were trivial compared with other leisure purpose. It suggested that the elders went out mainly for mental enjoyment instead of physical exercise [21, 22]. This was quite different from findings in some studies. Besides *Activity feature*, leisure purpose also directly affected *Leisure location*. This suggested the activity location and features were considered jointly once a desire for a leisure activity rose up. As for the effects of the leisure purpose on *Travel behavior*, instead of a direct impact as assumed in Fig. 1, it had only an indirect effect through activity features and location (see Table 4). In other words, how to travel and when to travel were secondary decision upon activity features and locations. This result was consistent with what we observed during survey conversations.

The environmental cognition was defined as people's mental action or process of acquiring knowledge about leisure activity environment factors through personal experience and other approaches. The results showed that the external incentive had significant direct effect only on activity locations and did not have statistically significant direct effect on activity features and travel behavior. In other words, the knowledge of leisure environment did not play any role in the decision of leisure activities directly but boosted leisure location choice once leisure activities were determined. The environmental cognition affected travel behavior only indirectly through leisure location choice (see Table 4).

Table 4 Standardized indirect and total effects on travel behavior

		Indirect effects				Total effects			
		Travel mode	Travel distance	Departure	Travel frequency	Travel mode	Travel distance	Departure	Travel frequency
Leisure purpose	Physical exercise	0.019	0.016	0.011	−0.005	0.019	0.016	0.011	−0.005
	Visiting friends	0.261	0.225	0.148	−0.071	0.261	0.225	0.148	−0.071
	Relaxation	0.427	0.341	0.242	−0.116	0.427	0.341	0.242	−0.116
	Hobbies	0.169	0.160	0.096	−0.046	0.169	0.160	0.096	−0.046
	Entertainment	0.192	0.176	0.109	−0.052	0.192	0.176	0.109	−0.052
	Killing time	0.247	0.207	0.140	−0.080	0.247	0.207	0.140	−0.080
Personal attributes	Age	−0.098	−0.082	−0.056	0.027	−0.290	−0.247	−0.165	0.079
	Family status	−0.041	−0.035	−0.023	0.011	−0.177	−0.152	−0.100	0.048
Environmental cognition	Beautiful environment	0.470	0.405	0.267	−0.128	0.470	0.405	0.267	−0.128
	Fresh air	0.401	0.345	0.228	−0.109	0.401	0.345	0.228	−0.109
	Cleanliness	0.334	0.287	0.190	−0.091	0.334	0.287	0.190	−0.091
	Activity facilities	0.037	0.032	0.021	−0.010	0.037	0.032	0.021	−0.010
	Easy access	0.281	0.242	0.159	−0.076	0.281	0.242	0.159	−0.076
	Proximity	0.476	0.410	0.270	−0.130	0.476	0.410	0.270	−0.130
Leisure location	Leisure site class	0.000	0.000	0.000	0.000	0.535	0.443	0.304	−0.146
	Leisure site size	0.000	0.000	0.000	0.000	0.567	0.469	0.322	−0.154
	Comprehensive environment	0.000	0.000	0.000	0.000	0.514	0.426	0.292	−0.140

Among five variables of personal attributes, *Income*, *Gender*, and *Education level* did not have significant direct impacts on any of activity features, location choice, and travel behavior. *Age* had negative direct impacts on *Activity feature*, and *travel behavior* but no direct effect on *Leisure location*. This was consistent with what we found during the survey and other empirical studies [23–25]. In general, older people had lower mobility and thus had shorter trips, stayed outside shorter, and usually were involved in fewer activities in a trip (see Table 4). Similarly, family status had negative impacts on activity features, travel behavior, and leisure locations. The elders living in larger families spent more time with family members and had fewer outdoor activities. They spent more time in nearby neighborhood green fields and parks rather than in farther and larger city parks.

One important factor in the model was the relationships between activity features, leisure locations, and travel behavior. Both *Activity feature* and *Leisure location* had positive direct impacts on *Travel behavior*. Larger leisure locations which were usually city parks were far from residential areas (i.e. longer travel distance) and less accessible by walking and biking, and were less frequently visited by the elders. Leisure activities with companions and with more contents generated more bus trips but the travel frequency was lower. *Activity feature* had significant direct impacts on leisure locations. That implied that the activities with more contents, with companions, and had longer duration more likely occur in locations larger in size, typically city parks. Contrary to the assumption, *Leisure location* did not have direct impacts on *Activity feature*. It could probably be interpreted as that the question of "what to do" preceded the question of where to do it. The missing path of impacts from *Leisure location* to *Activity feature* simplified the model from a non-recursive model to a recursive model and the assumed endogeneity did not support by the data. Further, the missing path implied *Environment cognition* had no impact on activity features indirectly. Thus, age, family status, and leisure purpose jointly determined the activity features. The external incentives showed no direct impacts on activity features at all (see Table 4).

In this model structure, the indirect effects of the variables on leisure trips were worth of attention due to their magnitudes in the total effects. The indirect effects relied on the existence of direct effects between the constructs. As shown in Fig. 2 and Table 4, leisure purpose did not have direct impacts on travel behavior but did have substantial indirect positive impacts on travel mode, distance, and departure time, and negative impacts on travel frequency through its impacts through *Activity feature* and *Activity location*. A higher score in leisure purpose implied that the elders tended to use more bike or bus as travel mode, go out in non-peak hours, have longer trip length, and travel less frequently.

The indirect effects of Age and Family status had the same signs with their direct effects. Their total effects were strongly boosted by indirect effects through activity features and locations. This finding further confirmed that the elder elders would less likely walk and make more short trips.

Although environment factors did not have direct impacts on travel behavior, they had strong indirect effects through their impacts on leisure activity locations. In terms

of the magnitude of total effects, the environmental factors had larger impacts than leisure purpose and personal attributes on travel behavior.

The above discussions were based on the impacts of a single construct or variable. Figure 2 and Table 4 showed that the elders' observed travel behavior was a consequence of joint impacts of many variables. Some impacts showed as direct effects while some showed as indirect effects. The advantage of SEM lied at its ability to differentiate and quantify the direct and indirect effects with clear paths. In addition, SEM had the ability to test the hypothesis on the relationships between some variables of interest. These abilities of SEM helped us understand how a variable affected the elders' travel behavior and how big the effects were.

5 Conclusion

In a fast aging society such as China, the provision of public services and transportation services to the elders through planning and policy tools is important. Understanding elders' leisure activity features and travel behavior is crucial for local governments to make smart land use and transportation planning policies. In this case study, through quantifying the direct, indirect, and total effects of endogenous and exogenous variables, we demonstrated the unique role of each variable in determining leisure activity contents, locations, travel mode, and travel distance. We found that the internal motivation, i.e. the leisure purpose, dominated elders' leisure activities and locations while environmental cognition only played a partial role in choosing leisure locations. In addition, we found that environmental cognition and leisure location had larger impacts than internal motivation in travel mode and distance.

In the literature and our case study, elders spent more than two hours on average everyday on activities outside the home. Their travel modes were mainly walk, bike, and bus. Neighborhood green fields and parks were used most often by elders due to their close proximity. This was especially true for the elder elders. For this reason, walk was the predominant travel mode. Although elders who are 60 years old or more in China have free transit passes, they take buses only if the leisure locations cannot be accessed by walk and bike. In China, the local governments have demonstrated great initiative to build large city parks and theme parks. These parks provide multiple entertainment services to city residents, but are usually far from residential areas and the elders who cannot easily travel long distances. What the elders need are small neighborhood parks within close proximity. Free transit passes help the elders access entertainment locations but do not help to satisfy daily leisurely needs. Further, increasing small community parks will reduce the demand for transit from the elders, and thus reduce the congestion within the bus and traffic congestion on the road. The unique needs of the elders should be taken into account in the city's redevelopment plan and long term urban land use plan.

Acknowledgements This research was supported by the National Natural Science Foundation of China (71661019).

References

1. World Health Organization (WHO) http://apps.who.int/iris/bitstream/10665/186463/1/9789240694811_eng.pdf. Accessed on 31 Oct 2015
2. China Population Statistics Yearbook. http://www.scio.gov.cn/xwfbh/xwbfbh/wqfbh/33978/34058/wz34060/Document/1462685/1462685.htm. Accessed on 10 July 2016
3. Wang L (2012) The research on elders leisure behavioral characteristics variation and leisure space organization. Chinese Academy of Sciences, Beijing
4. Zhang Z, Mao B, Liu M, Chen J, Guo J (2007) Analysis of travel characteristics of elders in Beijing. J Transp Syst Eng Inf Technol 7:11–20
5. Sun Y, Chen T, Han Y (2001) A study on leisure behavior of the aged in Beijing. Geogr Res 20:537–546
6. Long J, Wang S (2013) Serious leisure and happy life: a localization study based on Chinese senior group. Tour Trib 28:77–85
7. Guan X (2007) The research on outdoor leisure space of the elders in Guangzhou city: a case study in parks. Unpublished manuscript. Sun Yat-Sen University, Guangzhou
8. Crawford DW, Godbey G (1987) Reconceptualizing barriers to family leisure. Leis Sci 9:119–127
9. Crawford DW, Jackson EL, Godbey G (1991) A hierarchical model of leisure constraints. Leis Sci 13:309–320
10. Godbey G, Crawford DW, Shen XS (2010) Assessing hierarchical leisure constraints theory after two decades. J Leis Res 42:111–134
11. Wang W (2007) A study of leisure motivations and constraints for old people in Nanjing city. Unpublished manuscript. Nanjing Normal University, Nanjing
12. Rong P (2012) The research on people's leisure behavior and constraint factors—take the city of Kaifeng for example. Unpublished manuscript. Henan University, Kaifeng
13. Hayduk LA (1987) Structural equation modeling with LISREAL: essentials and advances. The John Hopkins University Press
14. Qiu H, Lin B (2009) The principle and application of structural equation modeling. China Light Industry Press, Beijing
15. Chen S (2010) Psychology of behavior. Hunan Normal University Press, Changsha
16. Dann GMS (1981) Tourist motivation an appraisal. Ann Tour Res 8:187–219
17. Goossens C (2000) Tourism information and pleasure motivation. Ann Tour Res 27:301–321
18. Jang S, Cai L (2002) Travel motivations and destination choice: a study of British outbound market. J Travel Tour Mark 13:111–133
19. Son JS, Mowen AJ, Kerstetter DL (2008) Testing alternative leisure constraint negotiation models: an extension of Hubbard and Mannell's study. Leis Sci 30:198–216
20. Zhao X, Guan H (2012) Modeling correlation of family holiday activities based on structural equation model. J Transp Syst Eng Inf Technol 12:135–14
21. Li M, Li M, Zhang M, Xie Y, Bian Y, Zhang M, Yang Z (2011) The effect of elders entertainment way on subjective well-being in Beijing. Chin J Gerontol 31:675–677
22. Li Z, Lu C (2004) Basic leisure science. Social Sciences Academic Press, Beijing
23. Agahi N, Parker MG (2005) Are today's older people more active than their predecessors? Participation in leisure-time activities in Sweden in 1992 and 2002. Ageing Soc 25:925–941
24. Dodge HH, Kita Y, Takechi H, Hayakawa T, Ganguili M, Ueshima H (2008) Healthy cognitive aging and leisure activities among the oldest old in Japan: Takashima study. JS Gerontol Ser A Biol Sci Med Sci 63:1193–1200
25. Horgas AL, Wilms HU, Baltes MM (1998) Daily life in very old age: everyday activities as expression of successful living. Gerontologist 38:556–568

Multi-scenarios Behavior Choice Model of Shared Parking in Residential Area

Jun Chen, Zexingjian Du and Jingheng Zheng

Abstract With the development of economics, the parking contradiction has become more serious in residential area. For optimization of the distribution of parking resource in residential area with limited land resources, the strategy of shared parking has been studied and adopted by more and more countries. The aim of this research is to determine the influence factors of selecting shared parking facilities in residential area and to increase the probability of choosing shared parking can increase with changing the characteristics of people or parking facilities. It will help to relieve parking contradiction and optimize parking resources in residential area. In order to acquire people's parking behavior characteristics, the paper designed questionnaire with three scenarios which contain shared parking and non-shared parking facilities. By questionnaire data introduced, the discrete choice modeling was set up to investigate the variability of probability of choosing shared parking across individual characteristics, socioeconomic attributes, trip and parking attributes, desire of accepting or providing shared parking and parking attributes in scenarios. For simulating the nonlinear effects of variables on the target variable more accurately, the BP neural network was used to filter redundant attributes. According to the model estimation results, charging had a significant impact on the selecting shared parking facilities. It found that people prefer selecting shared parking, which is same as the analysis of questionnaire. This indicates that the BP neural network is used to filter factors optimally. Finally, for appealing to people parking in shared parking facilities, it is suggested that the strategy of decreasing charging in shared parking facilities should be adopted. And the paper explored details areas for future research.

J. Chen (✉) · Z. Du
Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Nanjing, China
e-mail: chenjun@seu.edu.cn

J. Zheng
Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing, China

Z. Du
Transportation Management Department, Suzhou, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_7

Keywords Shared parking · BP neural network · Multi-scenarios · Parking behavior · Logit model

1 Introduction

With the development of economics, the parking contradiction has become more serious especially in residential area. For optimization of the distribution of parking resource in residential area with limited land resources, the strategy of shared parking is studied and adopted by more and more countries. For instance, when the shared parking between office land and hotel land was carried out by Urban Land Institute, the peak parking demand was reduced by 11 and 27% [1]. However, shared parking as a new parking management, there are many factors which affect people to select shared parking spots. So, knowing the characteristics of parking behavior in the residential area which has shared parking spots has great significance for optimization of parking resource.

Many researches confirmed that driver's personal attributes and travel characteristics have a great influence on his/her parking behavior (Waerden et al. 2008). When selecting parking spots, the drivers who were unfamiliar with surrounding environment cared less about walking time and security, but paid more attention to parking fees and other factors (Ruisong et al. 2009). In addition, parking behavior had difference between weekdays and weekend [2]. The sensitivity of person to parking fee in weekday is far higher than in weekend [3]. Pan [4] stated that the parking behavior could be divided into three types: commuter parking, non-commuter parking and residential parking.

The linear model was widely used to analyze the characteristics of parking behavior. Simićević [5] set up Logistic model based on intention preference survey to forecast parking behavior affected by factors in terms of parking charge and parking time. Yu [6] established discrete choice model which focused on off-street and on-street parking to study the behavior of selecting parking lots.

By summarizing, some weak points of current studies about parking behavior are as follows. Firstly, the current studies focused less on parking behavior in a certain shared parking environment. Secondly, the nonlinear effects of independent variables on dependent variables cannot be explained very well in Logit model which is often used to analyze parking behavior.

Unlike most of the previous studies, this study focuses on parking behavior in shared parking environment. The questionnaire was designed with three scenarios which contain shared parking spots. Moreover this study conducts discrete choice models to analyze influence mechanism of some factors on shared parking behavior.

The paper is structured as follows. Section 2 introduces the scenario of shared parking in residential area and describes the data of questionnaire. Section 3 presents the research methodology in terms of model specification, variable definition.

Section 4 discusses the estimation results of logit model and logit model with treatment of Back Propagation (BP) neural network. The last section concludes with the key findings and their implications for parking policies. Areas for future research are also identified.

2 Data Preparation

This paper focuses on shared parking which was carried out around the residential area, office land and commercial land. According to the difference of parking demand, three types of shared parking scenario were classified. Also, according to the difference between users and providers of spots, four types of shared parking modes were defined. The relative data of parking behavior was acquired by the questionnaire of situational hypotheses.

2.1 Shared Parking Mode in Residential Area

The parking space occupancy with land use in terms of residential area, office, hotel, market and hospital is presented in Fig. 1. It is clear that all kinds of land use have different parking peak time. Different land use attributes correspond to different parking requirements. For acquiring shared parking behavior with different travel purpose, three types of shared parking scenario were designed.

Scenario 1: Some nearby companies provide shared parking spots for residential area during peak parking demand period of residential area.

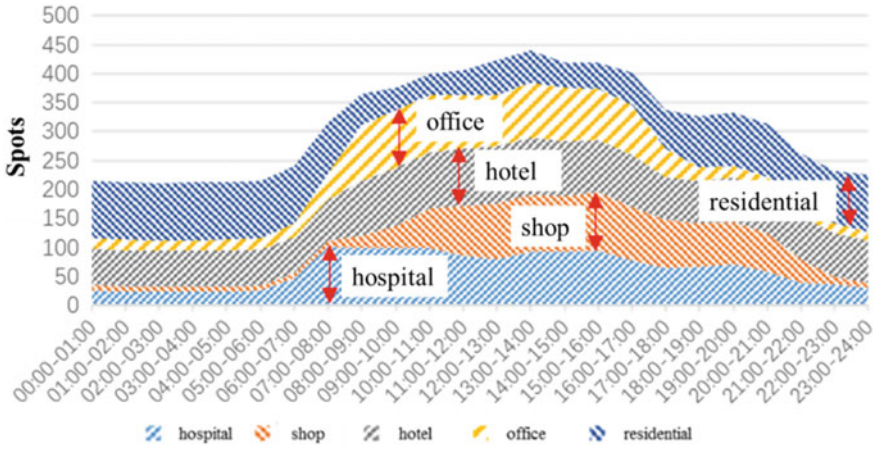


Fig. 1 The situation of parking occupancy in different land use

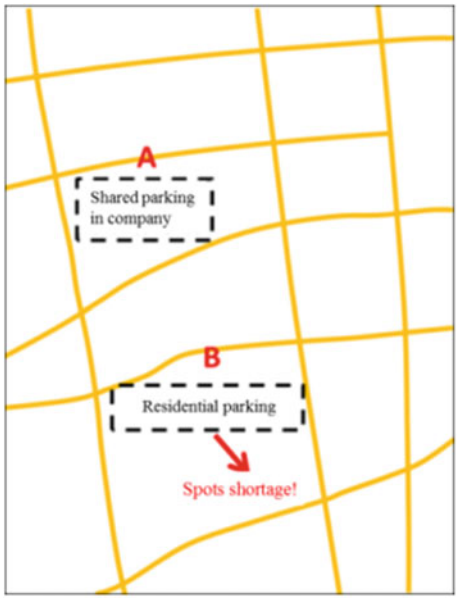


Fig. 2 Scenario 1

Scenario 2: Residential area provides shared parking spots for other land uses during flat parking demand period of residential area.
Scenario 3: Residential area provides shared parking spots for flexible parking demand such as shopping and travelling (Figs. 2, 3 and 4).

According to the difference between users and providers of spots, four types of shared parking modes were defined.

Mode 1: Using shared parking spots in residential area during flat parking demand period of residential area.
Mode 2: Providing shared parking spots in residential area during flat parking demand period of residential area.
Mode 3: Using shared parking spots in nearby companies during peak parking demand period of residential area.
Mode 4: Providing shared parking spots in nearby companies during peak parking demand period of residential area.

2.2 Traffic Investigation and Analysis Based on Scenario Assumption

According to definition in last section, this paper carried out stated preference survey to investigate parking people in Nanjing and Yangzhou. The questionnaire investigate Four attributes which are investigated by questionnaire are as follows.

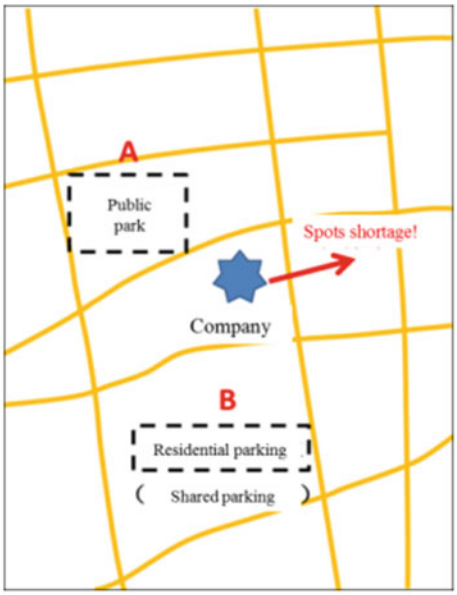


Fig. 3 Scenario 2

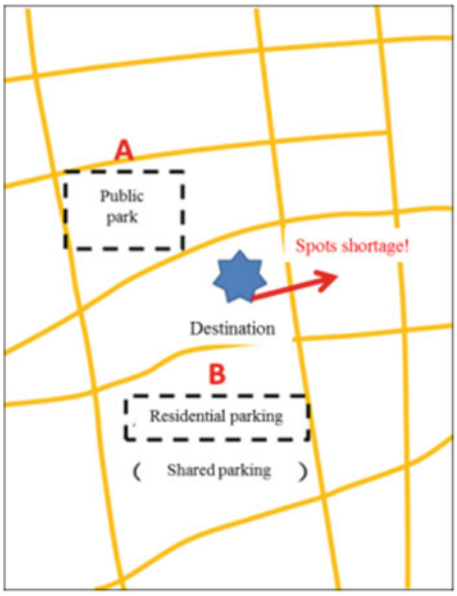


Fig. 4 Scenario 3

1. Social economic attributes (such as gender, age, driving experience and so on).
2. Daily trips and parking habits (such as time to work or go home, parking situation in home or company and so on).
3. Acceptability of four shared parking mode.
4. Parking characteristics in three shared parking scenario assumptions (Table 1).

Table 1 Descriptive Statistics for parking questionnaire

Attributes	Distribution	Percent	Attributes	Distribution	Percent
1. Social economic attributes					
Gender	Male	63.6	Driving age	Less 1 year	12.3
	Female	36.4		1–3 years	31.2
				3–5 years	21.5
Age	Less 20	0.5		5–10 years	21.2
	21–30	41.3		10 years and over	13.8
	31–40	33.8			
	41–50	20.5	Income	Less 50,000	16.0
	51–60	3.4		50,000–100,000	49.0
	60 and over	0.5		100,000 and over	35.0
2. Daily trips and parking habits					
Frequency of using car	1–3 a week	26.1	Parking occupancy in company	Saturated	46.8
	4–7 a week	36.0		Tight	40.2
	More than 7 a week	37.9		Enough	13.0
Familiarity of parking facilities	Unfamiliar	21.2			
	Familiar	49.3	Time to work	Before 7:00 am	23.9
	Very familiar	29.5		7:00–8:00 am	51.2
				8:00–9:00 am	22.5
Parking occupancy in residential area	Saturated	55.3		9:00–10:00 am	1.9
	Tight	36.3		After 10:00 am	0.5
	Enough	8.4			
Spot ownership in residential	Own	32.6			
	Rent	45.6	Time to go home	Before 16:00	2.7
	Not own	21.8		16:00–17:00	7.0
Spot ownership in company	Yes	30.0		17:00–18:00	32.8
				18:00–19:00	35.5
	No	70.0		After 19:00	22.0

3 Methodologies

3.1 Binary Logit Model in Several Scenarios

This paper studies the parking choice of residents with the existence of shared parking. Binary Logit Model (BL model) is the most popular discrete choice model [7]. By model analysis, the factors which affect residents selecting shared parking spots can be found. The random utility $U_{i,j}$ for each mode choice contains the systematic utility $V_{i,j}$ and the random residual $\varepsilon_{i,j}$. It is assumed that the random residuals independently and identically follow Gumbel distribution in a BL model. The random utility function of a choice j between select shared parking spot and not select it choice set J for a resident i is defined as

$$U_{i,j} = V_{i,j} + \varepsilon_{i,j} \tag{1}$$

Under the assumption of BL model, the probability of selecting parking choice j for the resident i can be expressed as

$$p_i(j) = \frac{\exp(V_{i,j})}{\sum_{k=1}^{J} \exp(V_{i,k})} \tag{2}$$

In mode utility functions, three kinds of attributes are considered [1, 8], that is, social economic attributes, daily trips and parking habits attributes and parking characteristics in three shared parking scenario assumptions. The definitions and denotation of detailed variables for these three kinds of attributes are illustrated in Tables 2, 3 and 4.

Because of the difference of parking purpose among three scenarios, the factors X22 and X23 are not put in scenario 1 and 2.

Table 2 Social economic attributes

Denotation	Definition	Detailed variable
X1	Sex	1 for male, 2 for female
X2	Age	1–6 represents six types of age
X3	Driving age	1–5 represents five types of driving age
X4	Income	1–3 represents 3 types of Income

Table 3 Daily trips and parking habits

Denotation	Definition	Detailed variable
X5	Driving frequency	Divided into 3 categories for 1–3
X6	Weekday parking time in the daytime	Divided into 3 categories for 1–3
X7	Weekend parking time in the daytime	Divided into 3 categories for 1–3
X8	Local parking facility familiarity	Divided into 3 categories for 1–3
X9	Spot condition in residential	Divided into 3 categories for 1–3
X10	Owning spot condition in residential	Divided into 3 categories for 1–3
X11	Spot condition in company	Divided into 3 categories for 1–3
X12	Owning spot condition in company	Divided into 3 categories for 1–3
X13	Time to work	Divided into 5 categories for 1–5
X14	Time to go home	Divided into 5 categories for 1–5
X15	Acceptability of mode 1	Divided into 3 categories for 1–3
X16	Acceptability of mode 2	Divided into 3 categories for 1–3
X17	Acceptability of mode 3	Divided into 3 categories for 1–3
X18	Acceptability of mode 4	Divided into 3 categories for 1–3

Table 4 Parking characteristics in three shared parking scenario assumptions

Denotation	Definition	Detailed variable
X19	Fixed charge difference between shared and non-shared spot	Divided into 4 categories for 1–4
X20	Rented charge difference between shared and non-shared spot	Divided into 4 categories for 1–4
X21	Walk distance difference between shared and non-shared spot	Divided into 4 categories for 1–4
X22	Travel time	Divided into 4 categories for 1–4
X23	Parking time	Divided into 4 categories for 1–4

3.2 Logit Model After Independent Variable Screening with Bp Neural Network

The drawback of the logit model is that its forward backward screening variable method sometimes fails to simulate the nonlinear effects on target variables [9]. So this paper adopted BP neural network filtering independent variable before setting up logit model.

Neural network is the most typical multilayer feed forward neural network algorithm with back propagation. The main structure of BP neural network includes input layer, hidden layer and output layer, as shown in Fig. 5. The input layer and output lay represent independent variable and dependent variable respectively. Furthermore,

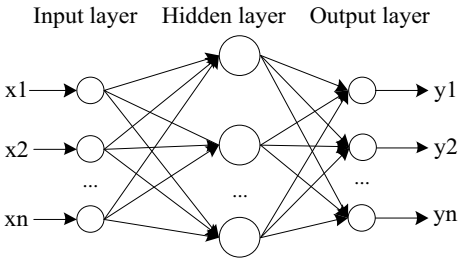


Fig. 5 Three layer neural network structure

the hidden layer denotes the influence mechanism of factors on target variables. Generally, the neural network system with three layer structure has the characteristics of approximating the objective function infinitely at arbitrary precision [10, 11]. Because of this, the BP neural network with three layers is adopted to carry out the analysis of variable significance screening.

The process of parking facilities selection is divided into three steps, as shown in Fig. 6.

The three sets of attributes mentioned above are conducted as input variables for BP neural network training samples and test samples. The number of nodes in input layer is determined to be 22 (21 in scenario 1). The nodes in output layer are whether residents select shared parking spots. The number of nodes in training samples and test samples are both 2. The basic neural network is structured as shown in Fig. 7.

4 Results and Discussion

4.1 Only Logit Model Result Analysis

The logit model is formulated with three set of attributes as independent variables in utility functions: social economic attributes, daily trips and parking habits attributes and parking characteristics in three shared parking scenario assumptions.

Table 5 present goodness of fit of the logit model predictions against the actual parking spots choices in three scenarios assumption. The goodness fit statistics denoting the percentage of observations correctly predicted indicate acceptable results. The parameter estimates for logit models are summarized in Table 6.

According to the estimation results of the logit model in scenario 1 with variables in Table 6, the respondents who are more familiar with parking facility information around their residential area are more inclined to select shared parking spots. Furthermore, the less difference of rented charge between shared and non-shared spots is. But, the difference of fixed charge is contrary to it. When the other variables remain unchanged, the walking distance reduced by 5 min, and the odds of shared spots is increased by 2.286 times. It reveals that the shorter the walk distance is, the higher probability selecting shared spots is.

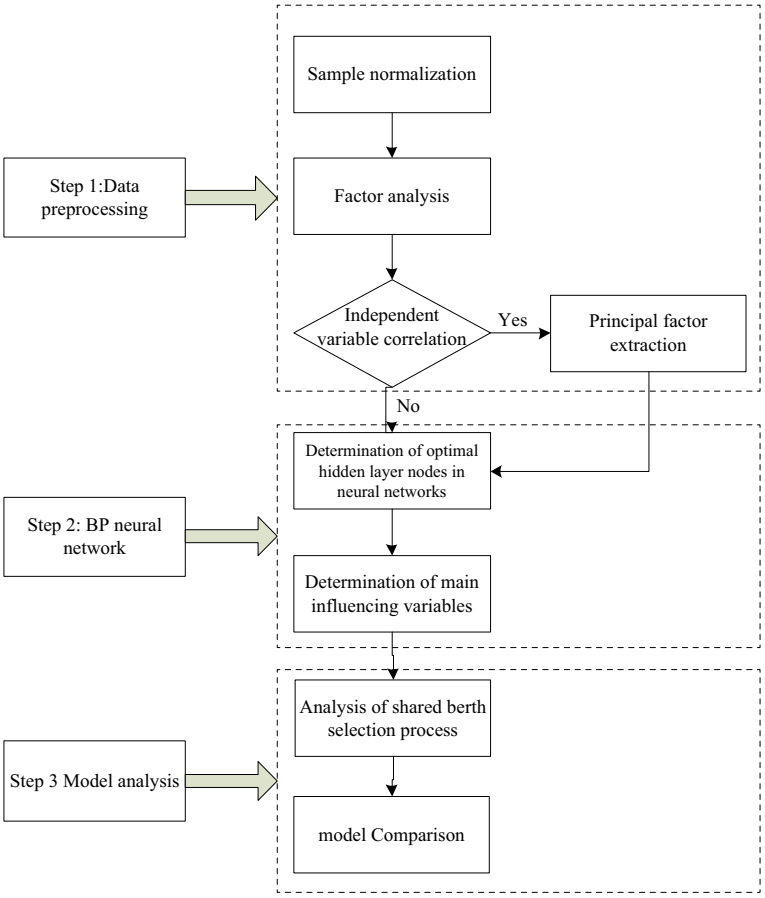


Fig. 6 Analysis of simulation shared spot selection process

Similarly, parking attributes in scenario 2 is used to formulate logit model in Table 6. It is found that people with high driving age prefer shared spots. From the results, it indicates that the staff whose company is with high occupancy rate of spots is more inclined to choose shared spots. And, the respondents who had no fixed spots had stronger willingness to choose shared spots. The influence of fixed and rented charge is similar to scenario 1.

From the results of logit model under scenario 3 in Table 6, people with high driving age prefer shared spots in scenario 3. Unlike scenario 2, the longer the vehicle parking time during the day, the more likely people choose shared parking space. However, this is the opposite of the weekend. The people with high-income is not inclined to select shared parking spots. When the parking time between 1 and 2 h, the odds of choosing shared spots is 0.730 time of parking time less than 1 h. It infers

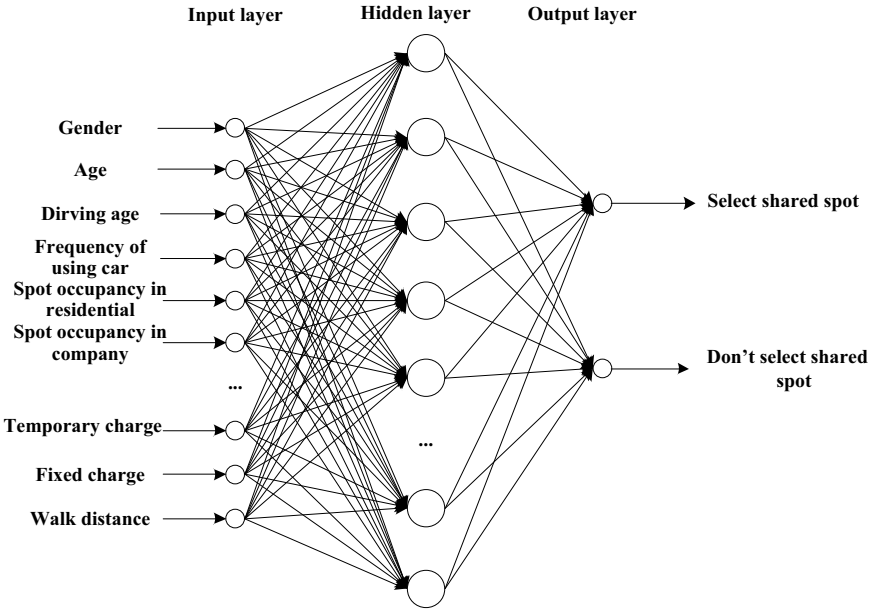


Fig. 7 Construction of neural network for independent variable selection process

Table 5 Goodness of fit for scenario assumption

	Goodness of fit (%)
Scenario 1	97.4
Scenario 2	93.5
Scenario 3	86.9

Table 6 Estimation Results of logit model

Variable	Scenario 1	Variable	Scenario 2	Variable	Scenario 3
X8	0.975 (0.000)	X3	0.393 (0.001)	X1	−1.064 (0.000)
X9	−1.140 (0.000)	X6	−0.472 (0.033)	X3	0.826 (0.000)
X14	−0.741 (0.000)	X8	−0.453 (0.021)	X4	−1.066 (0.000)
X17	−1.140 (0.000)	X11	−1.362 (0.000)	X5	0.349 (0.010)
X18	−0.706 (0.007)	X12	3.527 (0.000)	X6	0.534 (0.001)
X19	1.087 (0.000)	X15	−0.971 (0.000)	X7	−0.831 (0.000)
X20	−1.247 (0.000)	X16	−1.605 (0.000)	X8	−0.412 (0.005)
X21	0.827 (0.001)	X19	0.927 (0.000)	X20	0.555 (0.000)
Con	8.212 (0.000)	X20	−0.533 (0.000)	X21	−0.299 (0.017)
		Con	4.253 (0.000)	X23	−0.315 (0.001)
				Con	4.893 (0.000)

Note Parameter significance in parentheses

Table 7 The result of neural network screening (a)

Scenario type	Independent variable
Scenario 1	Price, Arrive, Fami
Scenario 2	Price-temp, Driage, Fami, Age, Park time
Scenario 3	Price-temp, Income

that people are more likely to select shared spots if they want to park in a short time. Similarly, the influence of rented charge is similar to scenario 1.

4.2 Logit Model Results After Bp Neural Network Filtering

The primary importance and standardization importance of each independent variable under three conditions, scenario 1, scenario 2 and scenario 3, are analyzed.

The variables of importance above 60% are taken as the independent variables of the modified shared spots selection model, and the results are shown in Table 7. The results of variables estimation are shown in Table 8.

In scenario 1, the parking demand is getting home. The number of spots decreases in the residential area after work time. It is clear that people would like to select shared spots in this situation. And if the charge of non-shared spots is more expensive than shared spots, people will be inclined to choose shared spots.

In scenario 2, the parking demand is getting working. When the rented charge of shared spots is lower than non-shared spots, yet people are not inclined to select

Table 8 The result of neural network screening (b)

Variable name	Scenario 1	Scenario 2	Scenario 3
Fami	0.414 (0.025)	−0.677 (0.000)	–
Driage	–	0.250 (0.077)	–
Age	–	−0.014 (0.944)	–
Income	–	–	−0.489 (0.000)
Arrive	−0.473 (0.003)	−0.741 (0.000)	–
Park time	–	0.054 (0.661)	–
Price-temp	–	−0.199 (0.155)	0.674 (0.000)
Price	0.457 (0.004)	–	–
Constant	1.854 (0.005)	3.921 (0.000)	0.586 (0.152)

Note Parameter significance in parentheses

shared spots. This paper suspects that the cause of this abnormality may be the relatively poor shared parking environment in residential areas. In scenario 2, Commuter parking time is very long. So, when the parking needs long time parking space, they would like to select shared spots.

In scenario 3, the parking demand is flexible, which is directly affected by money. It is indicated that high income people are not sensitive to the parking fees and the possibility of they selecting shared spots is decreasing.

The constant of model among three scenario assumption is positive. It is indicated that the respondents prefer to select shared parking spots under shared parking environment. The result is consistent with the descriptive analysis of questionnaire.

The coefficient of local parking familiarity is positive in scenario 1 but negative in scenario 2. This paper speculates that because of residents knows about the insufficient parking resources in residential area. In scenario 1, they need shared parking spots, but in scenario 2, the commuters are reluctant to use shared parking spaces in residential areas.

According to the results, this paper thinks that parking charge level is still a common significant factor in the three scenario assumption. For extension of shared parking spots and relieving city parking pressure, parking strategy which is aimed to reduce the charge of shared parking spots should be adopted.

5 Conclusions

This paper focused on parking behavior in the residential area and around office land and commercial land with the extension of shared parking spots. In order to acquire relative parking data, the questionnaire composed of different scenarios assumption was set up. According to different purpose, three kinds of scenarios which contain shared parking facilities were put up.

To analyze influence mechanism of some factors on shared parking behavior, the discrete choice models in different scenarios was conducted. For simulating the nonlinear effects of variables on the target variable more accurately, the BP neural network was used to filter redundant attributes. The main conclusions of this paper are as follows.

Based on variables screening in terms of BP neural network, the logit model was set up. And the main factors affecting the selection of shared berth were screened out.

According to the results of models, the respondents prefer to select shared parking spots under shared parking environment, which is consistent with the descriptive analysis of questionnaire. Furthermore, this paper agreed that parking charge level was still a common significant factor in the three scenario assumption.

Because of poor parking supply in residential area, the more residents are familiar with surrounding environment, the higher the percentage of choosing shared spots. And, the commuters are reluctant to use shared parking spaces in residential areas.

To popularize the shared parking, this paper suggested that the parking strategy which is aimed to reduce the charge of shared parking spots should be adopted.

The insufficiency of this article is that the further research is not conducted about the difference of coefficient of local parking familiarity between scenario 2 and scenario 3. It only speculates that it is concerned with parking conditions in residential areas. Future studies can be conducted to establish shared parking resource distribution model based on the characteristics of shared parking selection. Based on this model, the overall utilization efficiency of parking resources in residential areas and surrounding commercial and office areas could be enhanced.

Acknowledgements The research is supported by the Project of National Natural Science Foundation of China (Grant No. 51478111).

References

1. Xiu-cheng CYMG, Jiang-yu RAN (2010) The shared parking feasibility of appertaining parking facilities to the building in cities. Modern Urban Res 1:008
2. Liu H (2002) Study on parking behavior in big city's shopping center. Beijing University of Technology
3. Yao SY, Yan-Min LI (2008) Research on relationship between parking-charge and parking behavior in CBD. J Hebei Univ Technol
4. Pan C, Zhao S, Yao R (2012) Difference analysis of parking behavior based on trip purpose. J Transp Inf Saf 30(1):39–42
5. Simićević J, Vukanović S, Milosavljević N (2013) The effect of parking charges and time limit to car usage and parking behaviour. Transp Policy 30(3):125–131
6. Yu R, Yun M, Yang X (2009) Study on driver's parking location choice behavior considering drivers' information acquisition. In: Second international conference on intelligent computation technology and automation, vol 3, pp 764–770. IEEE Computer Society
7. Arora N (2010) Logit model. Wiley International Encyclopedia of Marketing
8. Dai L (2010) Research on the scale of comprehensive developed block's parking space base on 'shared parking analysis theory'. Southwest Jiaotong University
9. Anas A (1983) Discrete choice theory, information theory and the multinomial logit and gravity models. Transp Res Part B Methodol 17(1):13–23
10. Dia H, Panwai S (2011) Neural agent (neugent) models of driver behavior for supporting its simulations. Int J Intell Transp Syst Res 9(1):23–36
11. Stathakis D (2009) How many hidden layers and nodes? Int J Remote Sens 30(8):2133–2147
12. Pjhj WVD, Borgers A, Timmermans H (2008) Modeling parking choice behavior in enclosed business areas. Transp Res Board, Meeting

Pedestrian Arrival and Release Characteristics at Signalized Crosswalk

Xianmin Song, Di Liang, Lili Li, Qiujie Yang and Qiaowen Bai

Abstract The paper applied the shock wave theory to analyze the characteristics of pedestrian arrival and release, considering different related factors (i.e., crosswalk width, length, and signal timing). Then pedestrian release model with two constraints was established on the strength of above study. At last, correlation parameters were calibrated with the practical pedestrian data, and the lateral distance between pedestrians is shown as 0.8 m. The proposed model can be used in following two ways: the optimization of crosswalk width and the evaluation of crosswalk capacity to accommodate the demand. The result could offer some references to protect the safety of pedestrians and improve the crosswalk capacity.

Keywords Crosswalk capacity · Pedestrian arrival and release characteristics · Crosswalk width · Model application · Pedestrian safety

X. Song · D. Liang · Q. Bai
College of Transportation, Jilin University, No. 5988, Renmin Street, Changchun City, Jilin Province, People's Republic of China
e-mail: songxm@jlu.edu.cn

D. Liang
e-mail: 645758820@qq.com

Q. Bai
e-mail: 316545069@qq.com

L. Li (✉)
Transport Management Institute, Ministry of Transport of the People's Republic of China, Beijing, People's Republic of China
e-mail: 2441952937@qq.com

Q. Yang
TianJin Transportation Research Institute, No. 39, Pingshan Rd, Tianjin, People's Republic of China
e-mail: yangqiujie2009@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_8

1 Introduction

Pedestrian crossings are adopted to assurance the requirements for safety, convenience, speed and comfort of the pedestrians at most intersections and sections. Since the influence of signal control, the pedestrian release characteristics and release capacity at crosswalk are the most important factors that directly affect the number of pedestrians crossing the crosswalk safely. Pedestrian arriving characteristics, the width and length of the crosswalk and the signal timing plan are the key factors of affecting the pedestrian release characteristics and release capacity. The research on the pedestrians' arrival and release characteristics will help to determine the optimal value of crosswalk width and the capacity of pedestrian crossing, which also provide the basis for the design of crosswalk width and signal timing plan.

Most of the existing researches on the arrival and release characteristics of pedestrians are analyzed in a qualitative way. Zhang et al. [1] studied on the gathered group characteristics of pedestrian on sides of the crosswalk and the pedestrian queue dissipation characteristics at the beginning of releasing at the signalized intersection, and analyzed the causes of pedestrian group characteristics. However, his research was mainly based on the psychological characteristics of pedestrians which lacked quantitative analysis. Qu et al. [2] analyzed pedestrian arrival and release rule at signalized intersection using mathematical statistics method, then a pedestrian arrival distribution and pedestrian speed were modeled and calibrated. Nevertheless, this study was short of the detailed research on the release characteristics of pedestrians. A pedestrian dissipation time model was established after applying shock wave theory and stop-wave model by Aihajyaseen and Nakamura [3, 4]. The model was based on the analysis of Pedestrian arrival and release characteristics, but the dynamic changes of the pedestrian number and pedestrian interval distance in waiting area over time were not considered.

Studies on the design of crosswalk width mostly take pedestrian crossing process as the main stage. Such as, Aihajyaseen and Nakamura [3–5] analyzed the mutual influence between the number of pedestrians and crosswalk width which was constant. Feng et al. [6] observed and analyzed the walking characteristics of pedestrians; the reasonable walking clear width for the needs of pedestrians was obtained. Nevertheless, the study only considered only consider the spatial demand of pedestrians crossing the street, instead of studying the spatial requirement of the dissipation process when the pedestrians enter the crosswalk. Xu et al. [7] only gave a reference value of the crosswalk width. The level of service and pedestrian security basically is not considered in the design for the crosswalk width, and the effects of signal Settings are also rarely taken into account. The design value of crosswalk width mainly relies on the experience of the designer in practical application, which lacks scientific theoretical basis and practical guidance. Therefore, it is extraordinarily difficult to meet the security requirements when pedestrians cross the crosswalk during peak traffic.

The research on the capacity of crosswalk was initially based on the space occupancy theory [7], the pedestrian time occupancy theory [8], and the space-time occupancy theory [7]. With the gradual in-depth study, the factors such as the interaction of bi-directional pedestrians [9], signal timing and crosswalk width [10] were gradually taken into account. Above researches of crosswalk capacity were actually to find out the crosswalk containing capacity, ignoring the ability for pedestrians to enter the crosswalk within a specific period and the capacity to ensure the safety of pedestrians.

This paper combines the common factors in the previous studies and considers the changes of space gap when the pedestrians cross the street. The characteristics of pedestrians' arrival, accumulation and release are analyzed at signalized crosswalk to ensure safety of pedestrians and improve pedestrian crossing capacity at peak flow. A pedestrian release process model was established after considering the impact of related factors (i.e., crosswalk width, length and signal timing). On the basis of the parameter sensitivity analysis, this paper put forward the application of this model in two aspects: the optimal crosswalk width and crosswalk capacity of ensuring pedestrian safe. In addition, the proposed method can be regarded as the basis of crosswalk width and signal timing design.

2 Analysis and Modeling of Pedestrian Arrival and Release Characteristics

2.1 Analysis on the Pedestrian Arrival and Release Characteristics

Signal control of crosswalk can separate pedestrians from vehicles in time, but also has an accumulation phenomenon of pedestrians and vehicles in a certain time. Pedestrians arrive and gradually accumulate into a large crowd, the crowd in the waiting area will cross the street when the pedestrian light turns to green. The process can be seen as the form of wave from static to dynamic gradually. The arrival and release process of pedestrians is shown in Fig. 1.

2.2 Modeling of Pedestrian Arrival and Release Process

2.2.1 Modeling Methodology

After the signal light turns to green, pedestrians in the waiting area begin to enter the crosswalk, follow-up pedestrians will also arrive at the waiting area. When the pedestrian flow is high, the pedestrian crowd in the waiting area will be in a traffic

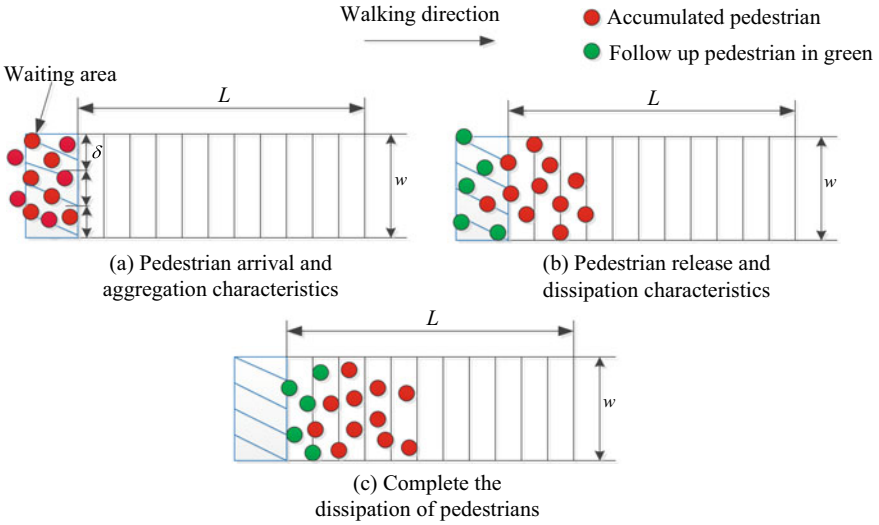


Fig. 1 The process of pedestrian arrival and releasing

waveform to dissipate to the crosswalk. The formation process of pedestrian crowd in the waiting area is illustrated in Fig. 2.

Therefore, the number of pedestrians walking into the waiting area minus the number of pedestrians entering crosswalk is the changing pedestrian numbers in waiting area. With the pedestrian green time going, dynamic changes of pedestrian numbers and pedestrian density in waiting area will also occur, thus Formula (1) can be obtained.

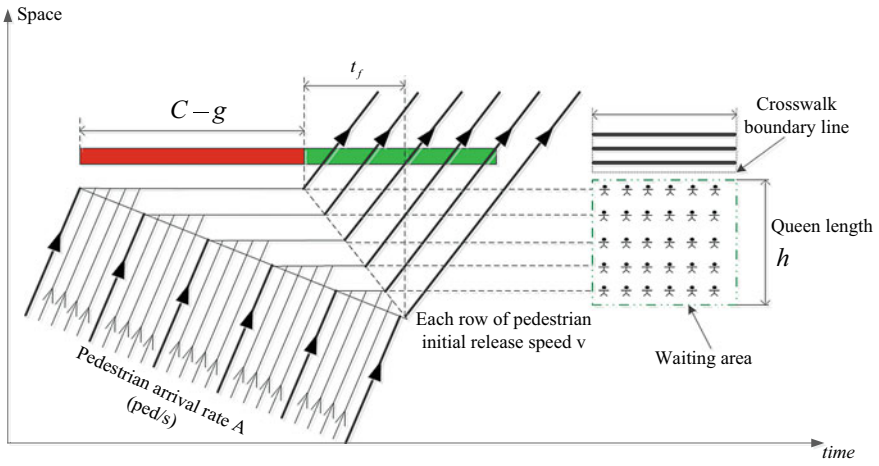


Fig. 2 A formation process of pedestrian crowd in the waiting area

$$A \times dt - v \times w \times k \times dt = d(k \times w \times h) \tag{1}$$

where, A is the pedestrians' average arrival rate (ped/s). v is the each row of pedestrians' instantaneous speed (m/s) by crossing the crosswalk boundary line. w is crosswalk width (m) (the width of waiting area). k is the pedestrians average density (ped/m^2) in waiting area, and assuming that regional density equals to the outflow density in pedestrian dissipation. h is pedestrian queue length (m), also equals to the length waiting area. And h is related to pedestrian numbers of the time begin to release and crosswalk width.

After solving the above formula and deforming it, Formula (2) can be obtained:

$$\frac{dk}{dt} + \frac{v}{h} \times k = \frac{A}{w \times h} \tag{2}$$

In any crosswalk, crosswalk width w is usually regarded as a certain value, pedestrian average arrival rate A and the instantaneous speed v of each row pedestrians crossing the boundary line can be determined through the investigation. h is a preset value of the length of waiting area (i.e. queue length), can also be treated as a constant value and independent of the duration of pedestrian green time. Therefore, the differential equation can be considered as a first order linear non-homogeneous differential equation.

The differential equation form is $\frac{dy}{dx} + p(x)y = q(x)$, and the general formula is $y = e^{-\int p(x)dx}\left[C_0 + \int q(x) \cdot e^{\int p(x)dx}dx\right]$. On the basis of the above general formula, according to Eq. (2), Eq. (3) can be gotten,

$$k(t) = \frac{A}{w \times v} + C_0 e^{-\frac{v}{h}t} \tag{3}$$

where, C_0 is a constant that should be determined. In order to calculate the value of C_0, we should first judge the initial state and the initial condition. According to the above description, the moment of pedestrian green can be regard as initial state of pedestrians releasing, so the initial conditions can be defined as $t_0 = 0$, $k_0 = p/(w \times h)$. t_0 is the initial time of pedestrian releasing (s), k_0 is pedestrian density in waiting area in the initial time (ped/m^2) and p is the number of pedestrians in waiting area, namely the accumulated pedestrian numbers during the period of pedestrian red (ped). In addition $p = A(C - g)$, C is the signal cycle length (s), g is pedestrian green time (s). The initial value of the Formula (3) is determined by Formula (4) as follows.

$$k(t) = \frac{A}{w \times v} + \left(k_0 - \frac{A}{w \times v}\right)e^{-\frac{v}{h}t} \tag{4}$$

The flow rate of pedestrian walking into the crosswalk is $dN(t) = v \times w \times k \times dt$, then both sides of Formula (4) are multiplied by $v \times w$, Eq. (5) can be calculated as follows

$$N(t) = A + (k_0 \times v \times w - A)e^{-\frac{v}{h}t} \tag{5}$$

With the pedestrian green time going, the accumulated pedestrians in waiting area are gradually release. In order to ensure the safety of pedestrians in crosswalk to cross the street, Formula (5) is integrated as follows:

$$\begin{aligned} N &= \int_0^{t_f} A dt + \int_0^{t_f} (k_0 \times v \times w - A)e^{-\frac{v}{h}t} dt \\ &= \int_0^{t_f} k_0 \times v \times w \times e^{-\frac{v}{h}t} dt + \int_0^{t_f} A\left(1 - e^{-\frac{v}{h}t}\right) dt \end{aligned} \tag{6}$$

where, N is the number of pedestrian crossing safety (ped). t_f is the termination time to guarantee crossing pedestrian safety (s), as shown in Fig. 2.

2.2.2 Constraints Conditions

(1) In Eq. (6), the first part is the number of pedestrians crossing crosswalk, who arrive in the red time, the second is the number of pedestrians crossing crosswalk safely who arrive in the period of pedestrian green time. The value of t_f in formula is described as follows

$$t_f = g - L/u_0 \tag{7}$$

where L is the crosswalk length (m), u_0 is pedestrian crossing speed in the late of pedestrian green (m/s).

(2) In order to assure pedestrian crossing safety, the arriving pedestrian number must be greater than or equal to the number of pedestrians can safely across the street. So

$$M \geq N, M = A\left(C - g + t_f\right) \tag{8}$$

3 Data Collection and Parameter Calibration

3.1 Data Collection

For evaluating the undetermined parameters, this paper applied video technology to collect six different signalized crosswalks' pedestrian data in Changchun City,

Table 1 Surveyed sites' characteristics

	Intersection (section) name	Geometrical characteristics	Survey time	Average one-way pedestrian flow (ped/h)
Site 1	Yatai Road section	8.0 m× 40.0 m	8:30–9:45	360
Site 2	Renmin Street-Chongqing Road	6.0 m× 42.0 m	11:30–12:30	700
Site 3	Hongqi Street-Gongnong Road	6.0 m× 26.0 m	15:30–17:00	840
Site 4	Weat gate of Jinlin university, Nanling campus	4.0 m× 25.0 m	11:30–12:00	170
Site 5	Tongzhi Street-Ziyou Road (east)	6.0 m× 25.0 m	15:40–16:00	750
Site 6	Tongzhi Street-Ziyou Road (north)	5.0 m× 15 m	15:40–16:00	330

China. The manual statistics method was adopted to obtain required traffic data and important parameters. The survey site features are shown in Table 1.

3.2 Parameter Calibration

(1) The instantaneous speed v of each row pedestrians

The moment when pedestrians walk through the crosswalk boundary line is defined as the starting state or the accumulation state that pedestrians get rid of the queue. By investigating pedestrian data of the above six sites (316 samples), the distribution of instantaneous velocity when pedestrian cross boundary line was obtained. Due to the concentration of speed distribution, and the speed nearby 1 m/s, so it is determined $v = 1\,\mathrm{m/s}$.

(2) Length of waiting area h

H is the preset length of waiting area, and it is related to pedestrian numbers in waiting area at initial time and the crosswalk width. It is assumed that pedestrians are arranged within a certain distance in the waiting area, there is

$$h = \frac{p}{(w/\delta)} \times \Delta t \tag{9}$$

where, δ is the lateral distance between pedestrians along the direction of crosswalk width (m), as shown in Fig. 1. Δt is the average time of each row pedestrians from the beginning to fully enter the crosswalk (s). According to 230 samples from the

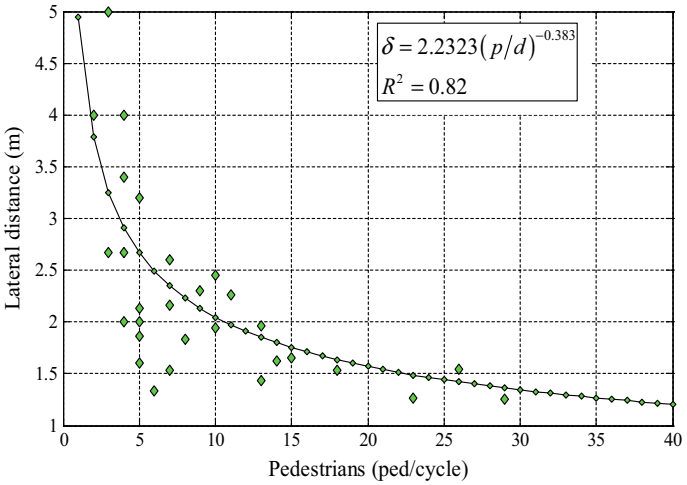


Fig. 3 Correction of pedestrian lateral distance δ

investigation sites 1, 2, 3, and calculating the average value, we can get $\Delta t = 1.2$ s and $RMSE = 0.37$.

The lateral distance δ of pedestrians refers to the basic form of the research given in Alhajyaseen [9], and then the coefficient corrected by the survey data of six sites. The data correction result is $\delta = 2.2323(p/d)^{-0.383}$, as shown in Fig. 3.

The parameter δ has physical meaning. with increasing of crossing pedestrian numbers, the lateral distance between pedestrians cannot be reduced infinitely. The limit value of the crowded lateral distance between pedestrians should be determined. According to the results from Fruin [11], the pedestrian static lateral space is 0.6 m and pedestrian required minimum standing lateral distance is 0.2 m. Therefore, the limit pedestrian lateral distance in waiting area is $\delta_{lim} = 0.8$ m.

(3) Pedestrian speed u_0 in the middle and later periods of pedestrian green time

Through the investigation, in the middle and later periods of pedestrian green time, the pedestrian average speed distribution is shown in Fig. 4.

Pedestrian speed increases obviously in the later stage of pedestrian signal green, but the frequency is low, so in a total perspective the speed distribution is uniform, and the mean value selected as the speed u_0 (m/s). Get $u_0 = 2.26$ (m/s).

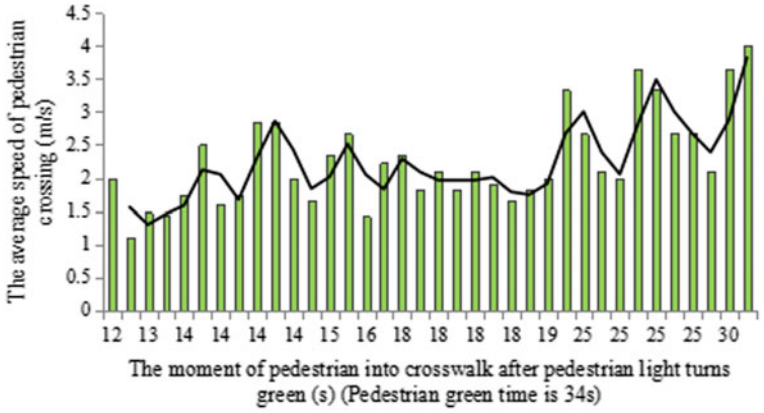


Fig. 4 The average crossing speed distribution in the later of pedestrian green

4 Sensitivity Analysis and Model Application

4.1 Sensitivity Analysis

The model not only describes the pedestrian arrival and release characteristics, but also calculates the number of pedestrians entering the crosswalk and safely crossing the street during pedestrian green time. The direct factor affecting the number of pedestrians crossing safely is pedestrian arrival rate. In addition, crosswalk length, width, and signal timing can also affect it. Figure 5 shows the impact of crosswalk length on the number of pedestrians safely crossing the street under the conditions of different pedestrian arrival rate. (Other parameters are the constant values).

As shown in Fig. 5, pedestrians will be more cautious with the extension of the crosswalk length if the signal timing is determined, so the number of pedestrians crossing safely will decrease. The number of pedestrians and the degree of congestion at waiting area in initial time are also different with different the pedestrian arrival rates. The pedestrian dissipation time will be longer with the increase of pedestrian arrival rate, and the influence degree of crosswalk length on crossing time is more.

Figure 6a shows the impact of crosswalk width on the number of pedestrians safely crossing the street under the conditions of different pedestrian arrival rates. As shown in Fig. 6a, if the signal timing plan is fixed, lateral distance of the waiting area and the whole space is larger when the crosswalk width is wider. Due to the limitation of crosswalk length and signal timing on the safe crossing time, the number of pedestrians can safely cross the street tends to reach a stable value.

Figure 6b shows the impact of crosswalk width on the number of pedestrians crossing the street safely under the conditions of different pedestrian green time. This figure also reflects the impact of the demand on crosswalk width of the pedestrian green time. If pedestrian green time is sufficient, and the pedestrian numbers in waiting area is definite, the less crosswalk width can satisfy the needs of pedestrian

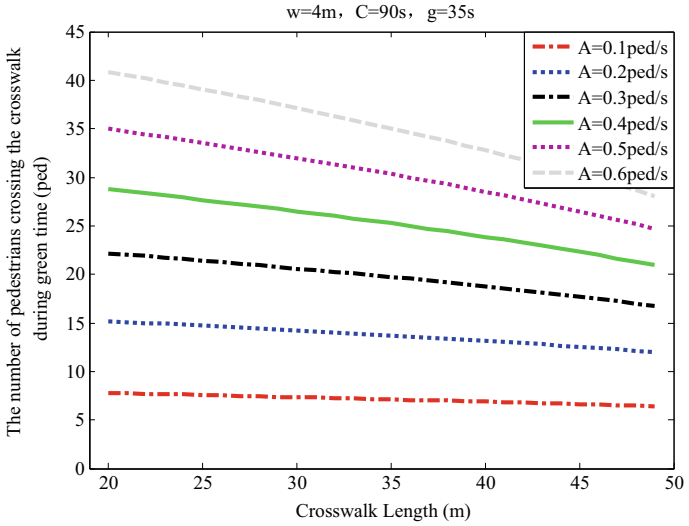


Fig. 5 Effect of crosswalk length on the number of pedestrians that crossing safely

safety to across the street, so the infinite increase of crosswalk width will only cause a waste of space resources. On the contrary, if pedestrian green time is not sufficient, the effect of crosswalk width on the number of pedestrians crossing safely is significant.

Figure 7 shows the impact of pedestrian green time on the number of pedestrians safely crossing the street under the conditions of different pedestrian arrival rate. The influence of green time on the number of pedestrians safely crossing the street is more significant, especially when a large number of pedestrians in the waiting area (the larger A). Moreover, the pedestrian green time increases to a certain extent, the effect of green time on the number of pedestrians crossing safely will reduce. In Fig. 8, the theoretical values are closest to the observed values when the walking speed is 1.14 m/s; at this point, there are 25 objective pedestrians on the crosswalk and a relatively comfortable lateral distance between them. Thus, the comfortable lateral distance is $\delta_{suit} = 1.29$ m.

4.2 Model Applications

Through the above analysis, the infinite increase of crosswalk width cannot continuously increase the number of pedestrians crossing the street safely. On the premise of improving the number of pedestrians crossing safely and crossing efficiency, the crosswalk width should be set reasonably. Taking the example of Fig. 6, the optimal crosswalk width is also different in the conditions of different pedestrian arrival rate and pedestrian green time. However, with increasing pedestrian arrival rate, the optimal crosswalk width (the width of crosswalk corresponding to the point where the

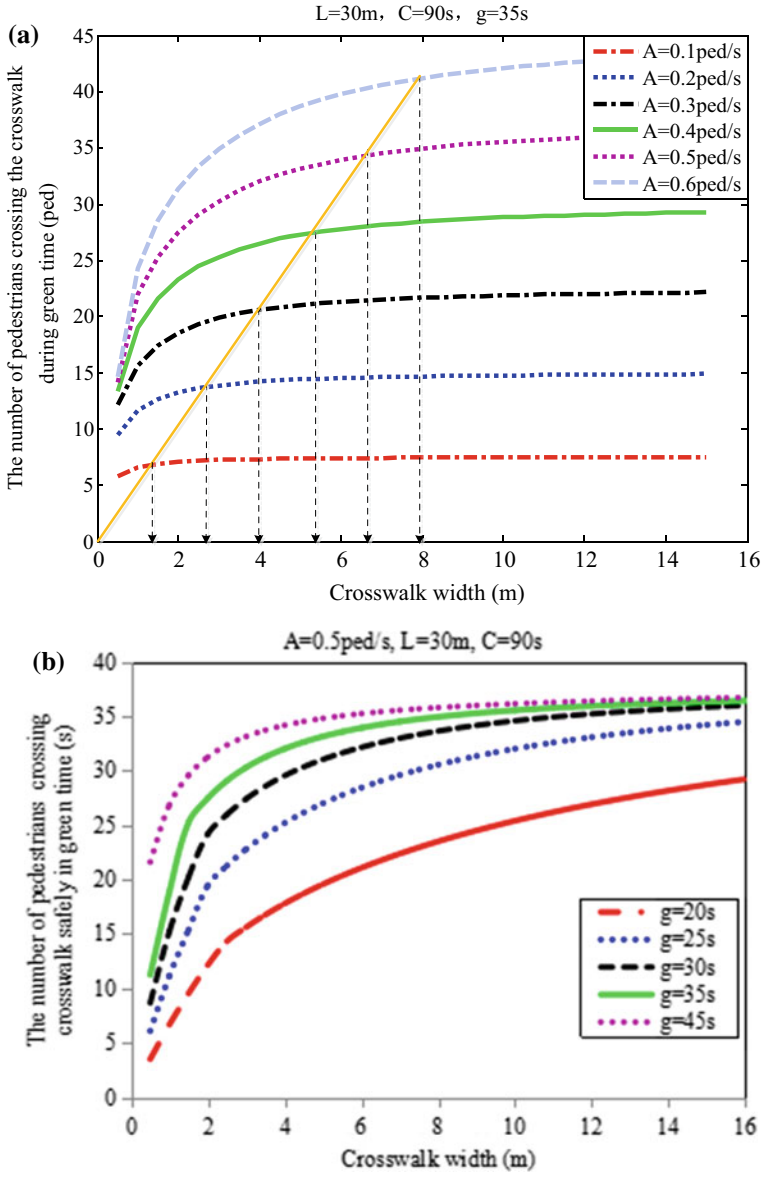


Fig. 6 **a** Effect of crosswalk width on the number of pedestrians that crossing safely. **b** Effect of crosswalk width on the number of pedestrians that crossing safely

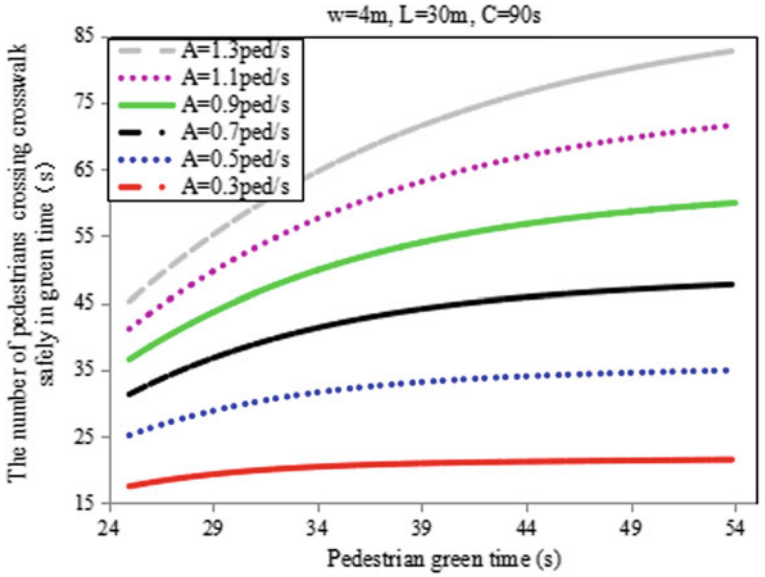


Fig. 7 Effect of pedestrian green time on the number of pedestrians that crossing safely

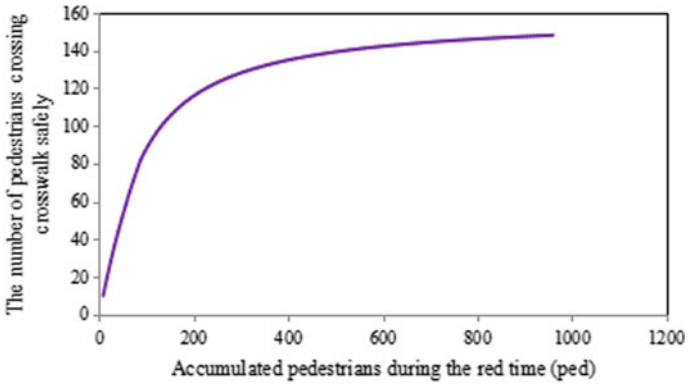


Fig. 8 The relationship between the accumulation pedestrian numbers and the number of pedestrians crossing safely

curvature tends to be stable) is gradually increased. Thus, the analysis of this model can be used to set the crosswalk width. In accordance with the above calculation method, we can get the optimal crosswalk width as shown in Table 2. Because the length of crosswalk is only related to the pedestrian crossing time, it is not related to the width value. So we only take L = 30 m as an example.

According to a comprehensive analysis of Figs. 5, 6a and b, in the case of pedestrian defined signal timing, pedestrian arrival rate can clearly calculate the number

Table 2 Optimal value of the crosswalk width

	L = 30 m	
	C = 90 s, g = 35 s	C = 120 s, g = 40 s
A = 0.1	1.8 m	1.6 m
A = 0.5	4.3 m	4.5 m
A = 1.0	9.2 m	9.8 m
A = 2.0	15.0 m	18.3 m
A = 5.0	19.0 m	33.0 m

of pedestrians in the waiting area of the initial state, and more the number of pedestrians arrive, more number of pedestrians cross the street. However, because of the limitation of the crosswalk width, the number of pedestrians crossing street safely cannot be indefinitely increased. When the pedestrian arrival rate is large enough, the number of pedestrians crossing the street safely will tends to a stable value, this value can be considered as the maximum number of pedestrians crossing the crosswalk per cycle.

As the crosswalk of its length is 30 m and width is 6 m (pedestrian arrival rate gradually increases) an example to illustrate, Fig. 8 shows the relationship between the accumulation pedestrian numbers during red light and the number of pedestrians crossing safely. Figure 8 shows the number of pedestrians in waiting area for more than 160 (the number of pedestrians can safely cross street is 108 ped, pedestrian arrival rate is 2 ped/s), increase rate of the number of pedestrians crossing safely becomes small. The number of pedestrians in waiting area for more than 528 (the number of pedestrians can safely cross street is 140 ped, pedestrian arrival rate is 6.5 ped/s), the number of pedestrians crossing safely tends to be stable. Therefore, the value of 140 is the maximum pedestrian numbers crossing the street per cycle of the signalized crosswalk.

The maximum number of pedestrians crossing the crosswalk per hour calculated as follows

$$C_q = \frac{3600}{C}\left[\int_0^{t_f} k_0 \times v \times w \times e^{-\frac{v}{h}t}dt + \int_0^{t_f} A_0\left(1 - e^{-\frac{v}{h}t}\right)dt\right] \tag{10}$$

where, C_q is the maximum number of pedestrians crossing the signalized crosswalk safely per hour (ped/h/w). A_0 is the pedestrian arrival rate when the number of pedestrians crossing the crosswalk safely tends to be stable (ped/s). Therefore A_0 is the key to calculate the maximum number of pedestrians crossing the signalized crosswalk safely. The above formula not only calculates the maximum number of pedestrians crossing the signalized crosswalk safely per hour, but also as a method to calculate pedestrian crosswalk capacity.

5 Conclusions

The paper applied the shock wave theory to analyze the pedestrian arrival and release characteristics. Considering the impact of crosswalk width, length, and signal timing and other factors, the pedestrian release model was established. Combined with the survey data, this paper puts forward the constraint conditions of protecting the safety of pedestrians and calibrates the lateral distance between pedestrians and its limit value, and gets the results that $\delta_{lim} = 0.8$ m. Through theoretical derivation, the model can calculate the number of pedestrians that crossing the street safely. The model includes two parts: the number of pedestrians arriving during the pedestrian red light can safely cross the street and the numbers of pedestrians can safety across the street in pedestrian green. Finally, this paper analyzes the influence of crosswalk width, length, and signal timing on pedestrian numbers that crossing street safely, the application of the model is proposed in two aspects: the optimal crosswalk width and crosswalk capacity that ensure the crossing pedestrian safety in specific conditions. In addition, the paper provides the crosswalk width recommended value in the different conditions of pedestrian arrival rate, crosswalk length and signal timing, then determines the signalized crosswalk capacity calculation formula.

The paper proposed the recommended value of crosswalk width can optimal meet the needs for pedestrians crossing street, but the crosswalk capacity is difficult to obtain through practical investigation. This paper only provides a method to calculate the crosswalk capacity, but not for the comparative analysis.

Acknowledgements This work was financially supported by the National Natural Science Foundation (51278220), (51278520), and Jilin major Project of Science and Technology Agency (20130206093SF).

References

1. Zhang H, Yu Q, Rong J, Wang D (2010) The research of bi-directional pedestrian platoon crossing characteristics at signalized intersection. In: 2010, 3rd international conference on power electronics and intelligent transportation system, vol 6, pp 136–141
2. Qu Z, Zhou L, Wang D (2004) Bicycle and pedestrians'arrival and departure characteristics at signalized intersection. J Highw Transp Res Dev 21(8):91–94
3. Alhajyaseen W, Nakamura H (2009) A methodology for modeling pedestrian platoon discharge and crossing times at signalized crosswalks. In: The 88th annual meeting of the transportation research board, Washington D.C., 11–15 Jan 2009
4. Alhajyaseen W, Nakamura H (2010) Estimating the minimum required width of signalized crosswalks considering bi-directional pedestrian flow and different age groups. J East Asia Soc Transp Stud 1(2):181–198
5. Alhajyaseen W, Nakamura H (2010) Design criteria for crosswalk width and position at signalized intersecyions. In: Proceedings of the 4th international symposium on highway geometric design, Transportation Research Board (TRB), Valencia, Spain
6. Feng S, Li Z, Zhang W (2008) Width of pavement in the city. J Haerbin Inst Technol 40(4):585–588

7. Xu J, Wen Y, Ni W (2011) Code for design of urban road traffic facility. Shanghai Construction and Transportation Commission. Ministry of Hou Human sing and Urban of People's Republic of China. Publication GB 50688–2011
8. Sisiopiku VP (2000) Probabilistic models for pedestrian capacity and delay at Roundabouts. Transportation Research Circular E-C018: 4th international symposium on highway capacity
9. Alhajyaseen W, Nakamura H, Asano M (2011) Effects of bi-directional pedestrian flow characteristics upon the capacity of signalized crosswalks. Procedia Soc Behav Sci 16:526–535
10. Li X et al (2012) Using cellular automata to investigate pedestrian conflicts with vehicles in crosswalk at signalized intersection. Discret Dyn Nat Soc
11. Fruin J (1971) Designing for pedestrians: a level of service concept. Metropolitan Association of Urban Designer and Environmental Planners Inc., New York, USA

Design and Operation Recommendations for Shared BRT Stops with No Overtaking

Jiao Ye, Jun Chen, Hua Bai and Dongping Li

Abstract In some BRT systems, certain sections of the operating infrastructure are shared with conventional buses. With increases in arrival rates and the occurrence of a phenomenon known as in-stop queueing, service quality and space utilization may decrease at shared BRT stops. This study analyzes the mixed operating characteristics at shared BRT stops and provides some suggestions for shared BRT stop design and operation to maintain service quality and space utilization. A wasted passageway calculation model is established to determine the space utilization. Relationships among service time, effective number of berths, stop capacity and expected number of wasted passageways are analyzed and corresponding design and operation recommendations are proposed.

Keywords Shared BRT stops · Stop capacity · In-stop queueing phenomenon · Wasted passageways

1 Introduction

Bus rapid transit (BRT) has emerged as a cost-effective alternative to metro systems and trams due to its high service quality and cheap fares. The main features that guarantee the efficiency and effectiveness of BRT systems include the use of high-capacity vehicles, high-frequency service, simple routes, dedicated lanes, off-vehicle ticketing and ITS applications [1]. Over the last few decades, an increasing number

J. Ye · J. Chen (✉)
School of Transportation, Southeast University, Nanjing, Jiangsu, China
e-mail: chenjun@seu.edu.cn

J. Chen
Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Nanjing, Jiangsu, China

H. Bai
China Design Group Co., Ltd., Nanjing, Jiangsu, China

D. Li
Shanghai Urban Construction Design and Research Institute (Group) Co., Ltd., Shanghai, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_9

of BRT systems have been built or are being built around the world to solve urban traffic congestion problems. In some BRT systems, certain sections of the operational infrastructure can be shared with conventional buses [2] because the BRT facilities may have redundant traffic capacities that cannot be utilized and thus become a barrier to transfers. To determine whether and how conventional buses can capitalize on BRT infrastructure, it should be premised that performance losses are not allowed. Stops are the pivotal component in determining the lowest capacity and overall performance along a bus route [3, 4]. Congestion may occur at shared BRT stops when arrival rates are beyond the maximum rate that buses can be discharged from a stop, subjecting the BRT system to queueing delays.

Unlike ordinary bus stops, most BRT stops have barriers between the platform and bus lane. The embedded paths in the barrier by which passengers board and alight the buses are called boarding and alighting passageways. Bus doors can only be opened when vehicles are berthed at the specified spaces that match the bus doors and passageways. Because the vehicle models and numbers of vehicle doors for general conventional buses and BRT buses differ, buses may arrive at a shared BRT stop when the number of remaining passageways is fewer than the number of doors. A special phenomenon, in-stop queueing, can be observed and is defined apart from the traditional queueing phenomenon. The remaining but useless passageways are wasted stop spaces under this condition. For high probabilities of the in-stop queueing phenomenon, stop space utilization will decrease.

Strategies for enhancing service quality at bus stops include designing the stop layouts, increasing stop capacities and controlling bus arrival rates [4–10]. The main purpose of these measures is to reduce the likelihood of bus queueing and time delays. Attention is continually paid to these three approaches, whereas the relationships among them are rarely considered.

Research has revealed that the capacity of a bus stop is determined by its size, especially the length of stopping bays or number of berths [11, 12]. Regarding linear loading areas with no overtaking, a downstream berth can provide less capacity than an upstream berth due to the effect of diminishing returns, whereby a stationary upstream bus may block another bus from entering the downstream berth. Loading efficiency factors were proposed, and the effective numbers of berths were calculated via simulation (see Table 1) [13]. To make full use of existing bus stops, berth optimized assignment methods have been proposed [14, 15] to reduce the number of idle berths due to the effects of diminishing returns. However, most studies on berth assignment have not considered that the original number of berths at a bus stop can differ because transit vehicles at the stop may have different sizes, especially at shared BRT stops with settled boarding and alighting passageways.

Table 1 Effective number of multiple linear berths at bus stops

Number of berths	1	2	3	4
Effective number of berths with no overtaking	1	1.75	2.45	2.65

To maintain the high service qualities of shared BRT systems and enhance the utilization of space at shared BRT stops, a wasted passageway calculation model to match stop design to frequency control is proposed in the present paper. The model connects shared BRT stop capacity, number of passageways, bus service times at a stop and ratio of BRT or conventional bus frequency. Data obtained at shared BRT stations in Changzhou, China, were used to calibrate the parameters and evaluate the model. Suggestions for shared BRT stop size (number of passageways) design and multi-line arrival rate coordinating are then offered.

The structure of the paper is as follows. The next section illustrates some performance parameters of shared BRT stops. Section 3 proposes a wasted passageway calculation model and uses the observational data to validate the model. In Sect. 4, the model is discussed, and some shared BRT stop design and operational recommendations are provided. Finally, the main conclusions are presented in the paper's last section.

2 Performance Parameters

2.1 Bus Dwell Time and Bus Service Time

The bus dwell time only considers the number of boarding and alighting passengers and the time used to open and close bus doors [16]. The first bus dwell time estimation model was proposed by Levinson [17]:

$$t_d = t_o + n \cdot \bar{t}_{ba} \tag{1}$$

where, t_d is the dwell time of a bus at a stop; t_o is the time required to open and close doors; n is the number of boarding and alighting passengers; and $\bar{t}_{ba}$ is the average time required by each passenger to board or alight a bus.

Bus dwell time is always used to estimate bus stop capacity [11]. A forward-positioned bus influences subsequent buses in the berthing stage and during the accelerating, decelerating and queueing processes [18]. The actual saturated headway therefore exceeds the bus dwell time. The above formulation should be revised to

$$t_s = t_d + t_m + t_q, \tag{2}$$

where, t_s is defined as the bus service time at a bus stop; t_m is the time taken by a bus entering and leaving the bus stop; and t_q is the time taken by a waiting bus to enter and leave the bus stop.

Assume that during the research period, the number of arrival records is m, and the service time of the i-th bus is t_s^i $(i = 1, 2, \ldots, m)$. The average bus service time during the research periods is then calculated using

$$\bar{t_s} = \frac{1}{m} \sum_{i=1}^{m} t_s^i. \tag{3}$$

2.2 Effective Number of Berths for Shared BRT Stops

The effective number of berths for traditional bus stops are exhibited in Table 1. However, the effective number of berths of shared BRT stops varies with the ratios of the two different kinds of vehicles. Let f_B denote the BRT frequency and f_T denote the conventional bus frequency; then

$$\begin{cases} \alpha = \dfrac{f_B}{f_B + f_T} \\ \beta = \dfrac{f_T}{f_B + f_T} \end{cases}. \tag{4}$$

Here, α is defined as the BRT frequency ratio, and β is the conventional bus frequency ratio. Moreover, α and β satisfy the following condition:

$$\alpha + \beta = 1. \tag{5}$$

It is obvious that the effective number of berths of a shared BRT stop is linearly related to the ratio of the two kinds of vehicles. The effective number of berths for shared BRT stops can be calculated by

$$N_e = \alpha n_{Be} + \beta n_{Te}, \tag{6}$$

where, N_e is the effective number of berths for the shared BRT stops; n_{Be} is the effective number of berths when the stop is only accessible to BRT; n_{Te} is the effective number of berths when the stop is only accessible to conventional buses. The values of n_{Be} and n_{Te} can be acquired via look-up from Table 1.

2.3 Capacity of Shared BRT Stops

The shared bus stop capacity model refers to the TCQSM (Transit Capacity & Quality of Service Manual) model [13]. The traffic blockage adjustment factors in the TCQSM model, for example lane type effect and influence at traffic signals, are assigned to 1 because the objective of this research is a single stop with no other interference. The bus stop capacity calculation model based on TCQSM can then be simplified to

$$C_a = \frac{\rho \cdot 3600}{\bar{t}_s} \cdot N_e. \tag{7}$$

Here, the definitions of $\bar{t}_s$ and N_e are the same as for Eqs. (3) and (6), and ρ is the expected saturability of the shared BRT stops.

2.4 Expected Wasted Passageways for Shared BRT Stops

A wasted passageway is defined as a remaining but useless passageway that appears during in-stop queueing. Regardless of stop size, there are only two scenarios associated with wasted passageways, which can be described as follows.

Scenario 1: One passageway remains. Whether the next arriving bus is a BRT or conventional bus, the remaining passageway will be wasted. **Scenario 2**: Two passageways remain, and the following bus is a BRT bus. If the following bus is a conventional bus, the stop is fully utilized, and no space is wasted. Each of these scenarios has a corresponding probability. Let $P_w(1)$ and $P_w(2)$ denotes the probabilities of scenarios 1 and 2, respectively, and the expected number of wasted passageway, $\bar{n}_w$, is then defined as

$$\bar{n}_w = 1 \times P_w(1) + 2 \times P_w(2). \tag{8}$$

The values of $P_w(1)$ and $P_w(2)$ are related to the total frequency of BRT and conventional buses, the proportion of arrival rates of different vehicles and the distribution of bus arrivals.

3 Calculation Model of Wasted Passageways

3.1 Model Description

Most BRT vehicles in service around the world have three doors, whereas conventional buses have two doors. Each door is used for both boarding and alighting. The service time for a single bus is determined by the largest number of passengers using two or three doors. Moreover, to study the in-stop queueing phenomenon, a proper stop size should be set for a shared bus stop. The number of passageways at a shared BRT stop depends on its location and the surrounding traffic conditions. Stops with six passageways are most popular in BRT systems, according to our observations. This research assumes that there are six passageways in the calculation model. In the discussion, the number of passageways is extended from three to nine.

As illustrated in the last section, there are two main wasted-passageway scenarios. As shown in Fig. 1a–d, four situations should be considered because their probabil-

(a) Scenario 1A (b) Scenario 1B

(c) Scenario 1C (d) Scenario 1D

(e) Scenario 2

Fig. 1 Scenarios of wasted passageways occurring at a shared bus stop with six passageways

ities differ from that of scenario 1. Scenario 2 only has one situation (see Fig. 1e). The description and probability of the five situations can be stated as follows.

Scenario 1: In this scenario, the first five passageways are occupied by one conventional bus and one BRT vehicle, regardless of whether the incoming bus is a BRT or a conventional bus. Different combinations in the arriving sequence result in the four specific scenario 1 situations.

Situation A: The first bus remaining at the stop is a conventional bus. The time gap t_{int}^{+} between the following BRT vehicle and the forward conventional bus is less than the service time of the conventional bus, t_{sT}. The next incoming bus is a conventional bus. The time gap t_{int}^{++} between the BRT vehicle and the following conventional bus is less than the service time of the front BRT vehicle, t_{sB}^{+}. The relevant probability of situation A is then calculated by

$$P_w^A(1) = \beta P_B(t_{\text{int}}^{+} < t_{sT}) P_T(t_{\text{int}}^{++} < t_{sB}^{+}). \tag{9}$$

Situation B: The first staying bus is a BRT vehicle, whereas the following two buses are conventional vehicles. The probability of situation B is

$$P_w^B(1) = \alpha P_T(t_{\text{int}}^{+} < t_{sB}) \times P_T(t_{\text{int}}^{++} < t_{sT}^{+}). \tag{10}$$

Situation C: The first bus remaining at the stop is a conventional vehicle, whereas the following two are BRT vehicles. The probability of situation C is

$$P_w^C(1) = \beta P_B(t_{\text{int}}^{+} < t_{sT}) \times P_B(t_{\text{int}}^{++} < t_{sB}^{+}). \tag{11}$$

Situation D: The first staying bus is a conventional bus, whereas the following vehicle is a conventional bus. The next in-stop queueing bus is a BRT vehicle. The probability of situation D is

$$P_w^D(1) = \alpha P_T(t_{\text{int}}^{+} < t_{sB}) \times P_B(t_{\text{int}}^{++} < t_{sT}^{+}). \tag{12}$$

The symbol '+' is used to represent a following bus, whereas the absence of the '+' symbol indicates the first staying bus at a shared BRT stop. Therefore, the probability of one wasted passageway at a shared BRT stop can be calculated as

$$P_w(1) = P_w^A(1) + P_w^B(1) + P_w^C(1) + P_w^D(1) \tag{13}$$

Scenario 2: In this scenario, the first four passageways are occupied by two conventional buses and the incoming bus is a BRT bus. The probability of two wasted passageways is calculated by

$$P_w(2) = \beta P_T(t_{\text{int}}^{+} < t_{sT}) \times P_B(t_{\text{int}}^{++} < t_{sT}^{+}). \tag{14}$$

The calculations of the probabilities of every scenario and situation are associated with the distribution of bus arrival times. Based on existing knowledge, vehicle arrival times at bus stops are usually assumed to follow a Poisson distribution [18], and the times between arrivals are exponentially distributed [19]. Let λ_χ denote the constant of an exponential distribution. The probability of the following bus arriving before the front bus leaves the stop is

$$P_\chi(t_{\text{int}}^{+} < t_s) = \int_0^{t_s} \lambda_\chi e^{-\lambda_\chi t} dt, \tag{15}$$

where, $\lambda_\chi = \alpha \cdot C_a^\chi / 3600$. Moreover, t_{int}^{+} is the time gap between the following and front buses, t_s is the service time of the front bus, and χ can be replaced by B or T, which respectively denote the service times and capacities, respectively, of BRT and conventional buses. The expected number of wasted passageways at a specific shared BRT stop can then be obtained when equations from (12) to (14) are substituted into Eq. (8).

3.2 Model Calibration

To calibrate the parameters in the calculation model, stop surveys were carried out over April 14–16, 2016. Six shared BRT stops with six passageways along BRT Line 1 and Line 2 in Changzhou, China, were chosen. The duration of the observations was from 7:00 to 18:00 every day. The recorded data include the arrival and service times of every berthed vehicle. A total of 1042 records containing peak hour data and non-peak hour data were obtained. The parameters differed little by service hour. For all the service times, the arrival moments were analyzed to justify the Poisson process assumption. The non-peak hour data passed the test, but the peak hour data

failed because of the bus bunching problem [20]. The model parameters were then calibrated using the non-peak-hour data. For convenience, average service time was used to replace the real service times of all the vehicles. The service saturation of the target shared BRT stop was assigned to 0.8, for example. The results of the model calibration are $\bar{t}_{sB} = 23.0$ s, $\bar{t}_{sT} = 21.2$ s, $\rho = 0.8$.

Moreover, $\bar{t}_s = \alpha \bar{t}_{sB} + \beta \bar{t}_{sT}$. Because $\bar{t}_s$ is defined as the mathematical average service time of all buses during the observation period, it is closely related to the value of α or β.

4 Discussion

4.1 Optimal Number of Passageways

In this section, seven groups of passageway numbers are considered. The numbers of passageways ranged from 3 to 9. The effective numbers of berths for BRT and conventional buses for different numbers of passageways can be obtained from Table 1. The values of n_{Be} and n_{Te} under conditions of four or six passageways were the same as those for conditions of five and seven passageways, respectively. There is no denying that the variations in the corresponding capacities of the shared BRT stops and the BRT ratio had the same tendencies, because capacity is linearly related to the effective number of berths (see Fig. 3). Figure 3 shows how the expected numbers of wasted passageways varied with BRT ratio for different numbers of passageways. For a specific BRT ratio and forecasted traffic demand, the optimal stop passageway number can be first determined from the corresponding lowest curve for the highest space utilization. The stop capacity is then examined to determine whether it matches the given traffic demand. If not, each sub-low curve is examined until the stop capacity satisfies the forecasted traffic demand.

For example, if the ratio of BRT vehicles is set to 0.3, and the forecasted traffic demand is 200 vehicles per hour, it can be seen from Fig. 3 that the corresponding lowest curve has five passageways. On checking the capacities in Fig. 2, stops with five passageways have capacities that exceed 200 vehicles per hour. Hence, the optimal number of passageways under these conditions is five. However, if the ratio of BRT vehicles is set to 0.7, and the forecasted traffic demand is 200 vehicles per hour, the best choice of passageway number is six rather than three because the lowest curve fails the capacity check.

Moreover, at many shared BRT stops, the ratios of BRT vehicles are not constant but vary with demand response. The optimal number of passageways cannot be determined using a fixed BRT ratio. The areas integrated under each curve in Fig. 4 were considered as an index to evaluate the stop usage efficiencies of different shared BRT stops. The results for different ranges of integration are displayed in Table 2. Most of the curves in Fig. 3 between 0.3 and 0.7 were lower than those outside that range. Not surprisingly, as seen in Table 2, when the range of BRT ratio was

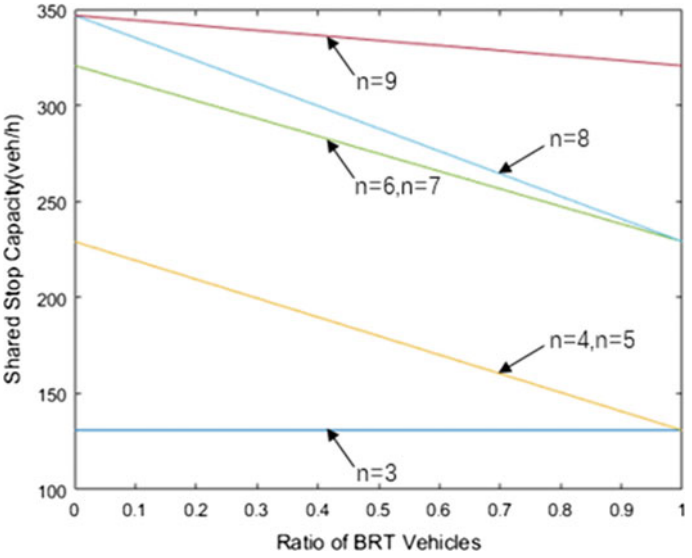


Fig. 2 The shared BRT stop capacity plotted against BRT ratio for different numbers of passageways

Table 2 The expected number of wasted passageways under different BRT ratio ranges

Number of passageways	3	4	5	6	7	8	9
Expected number of wasted passageways (0–1)	0.50	0.78	0.74	0.62	0.95	0.92	1.02
Expected number of wasted passageways (0.3–0.7)	0.20	0.36	0.26	0.36	0.43	0.41	0.57

0.3–0.7, each stop had high space utilizations with different passageways. Therefore, it is suggested that the BRT ratio be controlled between 0.3 and 0.7. After overall consideration, stops with six passageways are recommended when designing shared stop sizes.

4.2 Optimal Ratio of Arrival Rates

Figure 2 shows that the capacities of the shared BRT stops were inversely linearly related to the BRT ratios. The variation in the expected numbers of wasted passageways differed for different passageway numbers, as shown in Fig. 3. Optimal ratios of BRT arrival rates can be easily obtained from Fig. 3 given an acceptable expected number of wasted passageways. For example, if the acceptable expected number of wasted passageways is 0.6, the optimal ratio of BRT arrival rates should be between

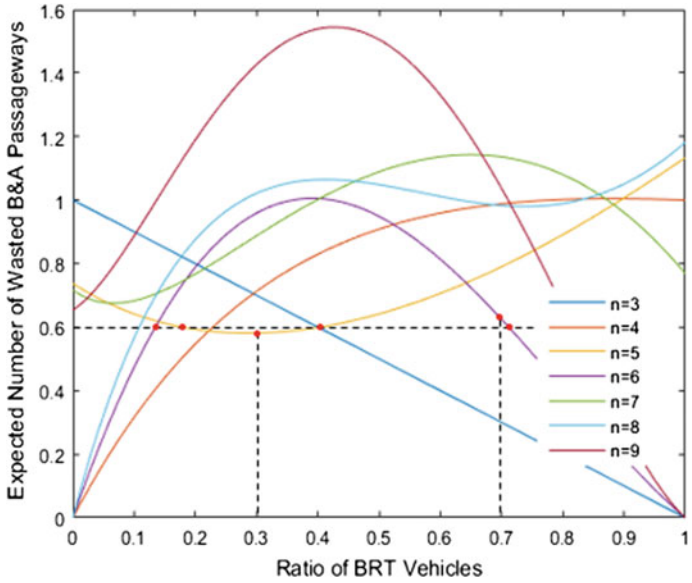


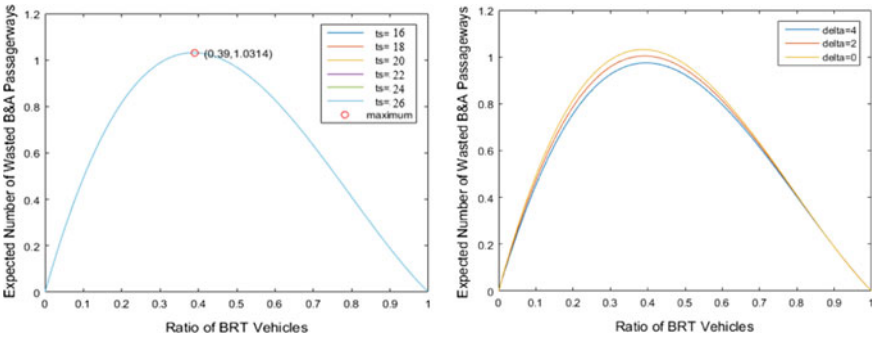
Fig. 3 The expected numbers of wasted passageways plotted against the BRT ratios for different passageway numbers

0.2 and 0.4 for 5 passageways. However, if there are 6 passageways, the optimal BRT arrival rate ratio should be less than 0.1 or greater than 0.7.

4.3 Service Time Consideration

Impacts of average service time and service time differences between BRT and conventional buses are examined in this section. The wasted passageway calculation model was employed to obtain the expected numbers of wasted passageways for different service times and service time differences. Based on the observations, the bus service times mainly ranged from 16 to 26 s. The service time differences never exceeded 4 s. In addition, the service times of BRT buses were usually longer than those of conventional buses.

The results are shown in Fig. 4. Six groups of numerical experiments were conducted to illustrate the impacts of different average service times (see Fig. 4a) when $\bar{t}_{sB} = \bar{t}_{sT}$. The expected number of wasted passageways was independent of the average service time but was relevant to the service time differences. Three curves for various service time differences are plotted in Fig. 4b. Delta is defined as the difference between service times of BRT vehicles and conventional buses. It is obvious that the time differences inversely impacted the expected numbers of wasted passageways. Because service time is determined by the number of boarding and alighting



(a) $\overline{t}_{sB} = \overline{t}_{sT}$ (b) Various Service Time Differences

Fig. 4 Plots of the expected number of wasted passageways against BRT ratio for different service times

passengers, the results indicate that the number of passengers per stop intending to board conventional buses should be controlled to be less than the number of passengers intending to board BRT buses. As a consequence, to maintain space utilization, it is best to not share BRT stops with conventional transit arteries because of the large passenger demands on conventional transit arteries.

5 Conclusions and Future Work

Good design and operation strategies for shared BRT stops are very important for maintaining high service quality and space utilization, and recommendations can be provided based on the research reported herein. The main performance parameters related to shared BRT stops are service time, stop capacity, effective number of berths and number of wasted passageways. This research carried out a systematic analysis of those parameters with the aim of identifying the optimal stop design and operation strategies. The relationships among the parameters were emphasized. The following conclusions were obtained from this research. (i) Five or six passageways are recommended at shared BRT stops because they showed the best comprehensive performance in terms of space utilization and service capacity. (ii) At a shared stop, 30% of the arriving buses should be BRT vehicles when there are five passageways, whereas no more than 40% of the arriving buses can be BRT vehicles when there are six passageways. (iii) Passenger demand for conventional buses should be controlled to be less than that for BRT vehicles. Branch lines of conventional buses are better choices for sharing.

Because this study assumed that the bus arrival process at a stop is subject to a Poisson distribution, the effects of bus bunching were neglected. Finally, it should be noted that capacity, as mentioned above, considers only vehicle capacity at a

shared BRT stop. Passenger capacity should be considered in following studies for comparison with vehicle capacity.

Acknowledgements This research was supported by National Natural Science Foundation Council of China under Project 51238008, Jiangsu Graduate Research Innovation Program KYCX18_0135, Shanghai Rising-Star Program (16QB1403000) and Shanghai Urban-Rural Development Transportation Talents Special Funds.

References

1. Levinson H et al (2002) Bus rapid transit: An overview. J Public Transp 5(2):1
2. Deng T, Nelson JD (2011) Recent developments in bus rapid transit: a review of the literature. Transp Rev 31(1):69–96
3. Gibson J, Baela I, Willumsen L (1989) Bus-stops, congestion and congested bus-stops. Traffic Eng Control 30(6):291–302
4. Fernandez R, Planzer R (2002) On the capacity of bus transit systems. Transp Rev 22(3):267–293
5. Barnett A (1974) On controlling randomness in transit operations. Transp Sci 8(2):102–116
6. Cats O et al (2010) *Mesoscopic modeling of bus public transportation*. Transp Res Rec J Transp Res Board 2188:9–18
7. Daganzo CF (2009) A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons. Transp Res Part B Methodol 43(10):913–921
8. Furth P, Rahbee A (2000) Optimal bus stop spacing through dynamic programming and geographic modeling. Transp Res Rec J Transp Res Board 1731:15–22
9. Estrada M, Ortigosa J, Robusté F (2011) Tandem bus stop capacity. In: TRB 90th annual meeting compendium of papers. 2011. Transportation Research Board of the National Academies Washington DC, USA
10. Al-Mudhaffar A, Nissan A, Bang K-L (2016) Bus stop and bus terminal capacity. Transp Res Procedia 14:1762–1771
11. Gu W et al (2011) On the capacity of isolated, curbside bus stops. Transp Res Part B Methodol 45(4):714–723
12. Manual HC (2010) Highway capacity manual. Washington, DC
13. Kittelson & Associates I et al (2013) Transit capacity and quality of service manual. TCRP. Transportation Research Board: Washington, D.C., pp 77–80
14. Wu X et al (2011) Berth assignment planning for multi-berth bus stops. In: 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC). IEEE, New York
15. Lu L et al (2010) Optimal design of bus stops that are shared by multiple lines of buses. In: 2010 13th International IEEE Conference on Intelligent Transportation Systems (ITSC). IEEE, New York
16. Meng Q, Qu X (2013) Bus dwell time estimation at bus bays: A probabilistic approach. Transp Res Part C Emerg Technol 36:61–71
17. Levinson HS (1983) Analyzing transit travel time performance
18. Bian B et al (2015) Bus service time estimation model for a curbside bus stop. Transp Res Part C Emerg Technol 57:103–121
19. Chriqui C, Robillard P (1975) *Common bus lines*. Transportation Science 9(2):115–121
20. Newell GF, Potts RB (1964) Maintaining a bus schedule. In: Australian Road Research Board (ARRB) conference, 2nd, 1964, Melbourne

Sensitivity of Simulated Conflicts to VISSIM Driver Behavior Parameter Modification

Qi-yu Liang, Qian Wan, Lu Bai, Hao Yu, Liu-xuan Lv and Dong-ping Li

Abstract The impact of driver behavior parameters in a widely used microscopic simulation package, i.e., VISSIM, on the simulated traffic conflict counts is presented in this study. More specifically, 31 driving behavior-related parameters is investigated with different parameter value. By modifying the parameter value, VISSIM produces different vehicle trajectory files, and a surrogate safety assessment model (SSAM) package is applied to output the conflicts. It shows that, in the present model, the numbers of parameters sensitive to the count of total conflict, rear-end conflict and lane-change conflict were 10, 12, and 4, respectively. The study partly demonstrated the impacts of different human behavior parameters provided in the VISSIM on the

Q. Liang · Q. Wan · H. Yu · L. Lv
Hualan Design and Consulting Group, 39 Hua Dong Road, Nanning 530011, China
e-mail: liangqiyu@126.com

Q. Wan
e-mail: 1564656392@qq.com

L. Lv
e-mail: 447369458@qq.co

Q. Liang · Q. Wan
Guilin University of Electronic Technology, 1 Jin Ji Road, Guilin 541004, China

L. Bai · H. Yu (✉)
School of Transportation, Southeast University, 2 Sipailou, Nanjing 210096, China
e-mail: seudarwin@gmail.com

L. Bai
e-mail: xinyuesther@126.com

H. Yu
Department of Civil Environmental Engineering, 2540 Dole St., Honolulu, HI 96822, USA

D. Li
Shanghai Urban Construction Design and Research Institute (Group) CO., LTD., 3447 Dong Fang Road, Shanghai 200125, China
e-mail: lidongping@sucdri.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_10

simulated conflicts. Future studies were to be conducted to sensitive analysis of multi-parameters and the evaluation countermeasures for the impacts of these parameters on simulated conflicts.

Keywords Sensitivity · Driver behavior · Traffic conflict · Simulation · SSAM

1 Introduction

Microscopic traffic simulation models are universally applied in different fields of transportation engineering, such as development and evaluation of transit signal logic, evaluation and optimization of traffic operations, analysis of traffic flow, and the evaluation and comparison of the safety assessment of design alternatives [1]. The Surrogate Safety Assessment Model (SSAM) is proposed by a research group in SIEMENS to identify conflicts from vehicle trajectory files generated by popular microscopic traffic simulation models, such as VISSIM, PARAMICS, AIMSUN, and TEXAS [2]. And by analyzing vehicle trajectories, the SSAM is able to export the conflict information including the counts, the types, and the locations. Simulated conflicts were classified into three types, which were rear-end conflict, lane-change conflict, and crossing conflict, by conflict angels.

By evaluating the relationships between simulated conflicts and crashes, several studies have been conducted to evaluate the validity of using simulated conflicts for safety assessment [3–7]. Even though these studies have illustrated the potential relationships between simulated conflicts and crashes, controversies are continuing regarding the validity of using simulated conflicts in safety assessment. Characteristics and behavior of individual vehicles are of great importance in contributing to the occurrence of simulated conflicts. One of the major concerns is that the impacts of driver behavior parameters on the occurrence of simulated conflicts are still unknown.

In the authors' previous studies, conflicts generated by VISSIM simulation model were compared to the field measured traffic conflicts at freeway merging areas [6]. In this study, a two-stage calibration and validation procedure was proposed. And it claimed that the consistency between the simulated and the observed traffic conflicts is improved by calibrating VISSIM simulation models and adjusting the threshold values used for defining simulated conflicts in SSAM. Following our previous study, the primary objective of this work is to explore the impacts of the driver behavior parameters in VISSIM model on simulated traffic conflicts. More specifically, the research objective of this study includes the following two tasks: (a) to develop VISSIM simulation models which take into consideration of the driver behavior parameters at freeway merging area and to calibrate the simulation model using field collected traffic conflicts; (b) to identify the impacts of various behavior parameters on traffic conflicts.

2 Field Data Collection

In this study, a set of VISSIM models were built based on a freeway merging section in Nanjing area in China. Two types of data were collected: (a) geometric data, such as the lane widths, length of acceleration lanes were obtained using a measuring wheel; (b) traffic data, such as traffic volumes and conflicts, was collected by means of video tapes. A total of 88 h of traffic video data was recorded at the selected sites. And vehicle speed records at freeway mainlines, and ramps were measured directly in the field using a radar gun.

Field data collection was conducted on weekday under normal weather conditions.

The recorded videos were later reviewed in the laboratory for obtaining traffic volumes and traffic conflicts. A trained graduate student was asked to review all the videos to ensure that consistent criteria were applied for identifying conflicts. Generally speaking, traffic conflicts were identified by vehicles' braking, swerving, and noticeable deceleration. Once a conflict was identified, the following information was collected: (a) the time of conflict when the first vehicle took an evasive action to avoid a collision; (b) the distance between conflicting vehicles and the conflict point; (c) the angle of conflicting vehicles when the evasive action was taken; (d) the speed of each conflicting vehicle. The traffic conflicts observed in the field were classified as two types, including 2563 rear-end conflicts and 1294 lane-change conflicts.

Collected data was divided into two data sets, i.e., the calibrating and the validating dataset. By applying the two-stage calibrating and validating procedure proposed in Fan et al. [5], the VISSIM models built in this study were well calibrated. Corresponding parameter settings were used as the default value in the following analyzing process.

3 Experiment Design

In order to capture individual behaviors, 31 driver behavior related parameters were employed in the proposed VISSIM models. These parameters could be divided into three groups, which were "following", "lane change", and "lateral", as shown in Table 1. A detailed description for each parameter was given when referred. As shown in Table 1, among the 31 parameters, there were 5 dummy variables, which would be evaluated at two different values, and 26 continuous variables, which would be evaluated at three level of magnitudes (LOM): low, medium, and high. As for the selection of different LOMs, the common ranges of the parameters were considered [8]. It should be emphasized that the experiments were not intended to prove a causal relationship between the selected parameters and the counts of simulated traffic conflict. But it would be viewed as a qualitative perspective as to the safety performance of the simulation at freeway merging areas when the driving behavior parameters modified.

Table 1 Value set for various driver behavior related parameters

Parameter name	Default value	Level of magnitude		
		Low	Medium	High
FOLLOW behavior				
Look ahead distance Min	0	0	50	100
Look ahead distance Max	250	200	300	500
Look ahead distance observed-vehicle	2	1	3	5
Temporary lack of attention duration	0	1	2	4
Temporary lack of attention probability	0	1	3	5
Wiedemann-99-CC0	1.5	0.6	1.2	3
Wiedemann-99-CC1	0.9	0.5	1	1.5
Wiedemann-99-CC2	4	1.5	3	6
Wiedemann-99-CC3	−8	−4	−10	−15
Wiedemann-99-CC4	−0.35	−0.1	−0.5	−2
Wiedemann-99-CC5	0.35	0.1	0.5	2
Wiedemann-99-CC6	11.44	2	8	20
Wiedemann-99-CC7	0.25	0.15	0.3	0.5
Wiedemann-99-CC8	3.5	2	3.6	5
Wiedemann-99-CC9	1.5	0.6	1.2	2,4
LANE CHANGE behavior				
General behavior	Free lane selection	Right side rule		
Own maximum deceleration	−4	−3	−4.2	−5
Own −1 m/s^2/distance	200	160	210	250
Own accepted deceleration	−1	−0.6	−1.2	−1.8
Trailing maximum deceleration	−3	−1.2	−2.5	−3.5
Trailing −1 m/s^2/distance	200	160	210	250
Trailing accepted deceleration	−0.5	−0.3	−0.6	−1.2
Waiting time before diffusion	60	40	70	120
To slower lane if collision above	0	0.5	1.2	2
Safety distance reduction factor	0.1	0.01	0.2	0.6
Maximum breaking deceleration	−9	−5	−8	−10
Min headway	0.5	0.3	0.8	1.5
LATERAL behavior				
Desired position at free flow	Middle of lane	Any	Right	
Observed vehicle on next lanes	No	Yes		
Diamond shaped queuing	No	Yes		
Over take on same lane	No	Right	Left	Both

Coping with the random fluctuations, the authors conducted multiple simulation-runs for each parameter value. The number of simulation-runs satisfied the minimum number constraint estimated by the sample mean and the standard deviation [9]:

$$N \geq (t_{\alpha/2} * \frac{\sigma}{\mu * \varepsilon}) \tag{1}$$

where μ and σ were the mean and standard deviation of the corresponding variable, respectively; ε was the allowable error specified as a fraction of the mean; and $t_{\alpha/2}$ was the critical value of the t-distribution at $1 - \alpha$ level of confidence.

Student's t-tests were conducted to identify if the difference in simulated conflict counts between different parameter value and the default value was statistically significant. The null hypothesis of the tests was that the average counts of simulated traffic conflict recognized by SSAM with different parameter values were the same. The null hypothesis would be rejected if [10]:

$$t = \frac{\left|C_{V2,avg} - C_{V1,avg}\right|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \geq t_{\alpha/2,df} \tag{2}$$

where α was the level of significance; $C_{V1,avg}$ and $C_{V2,avg}$ were the average counts of traffic conflict recognized by SSAM; S_1 and S_2 were the sample standard deviations; n_1, n_2 were numbers of simulation runs conducted with each parameter value; $t_{\alpha/2,df}$ was $100(1 - \alpha/2)$ percentile of the t distribution with the degree of freedom df given by

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}} \tag{3}$$

4 Result Analysis

The number of simulation runs for each parameter value was 250, which satisfied the requirements introduced by Eq. (1). The results of the Student's t-tests shown that 13 out of 31 parameters rejected the null hypothesis of no significant difference at 95% level of confidence. More specifically, 10 parameters were sensitive to the count of total traffic conflict, 12 parameters were sensitive to the count of rear-end conflict, and 4 parameters were sensitive to the count of lane-change conflict. All the sensitive parameters were shown in Table 2 together with the results of Student t-test. And graphical descriptions of multiple simulation runs were summarized in

Table 2 T-test results for parameters sensitive

No	Parameter name	Parameter value	T[a]	*p*-value R[b]	L[c]
1	Look ahead distance Max Default = 250	200	0.44	0.28	
		300	0.02	<0.01	
		500	0.37	0.61	
2	Look ahead distance observed vehicle Default = 2	1	<0.01	<0.01	
		3	0.13	0.98	
		5	0.32	0.01	
3	Temporary lack of attention duration Default = 0	1	0.32	0.3	0.07
		2	0.13	0.85	0.04
		4	0.01	0.54	<0.01
4	Temporary lack of attention probability Default = 0	1	0.32	0.3	0.07
		3	0.01	0.2	0.01
		5	0.02	0.02	0.07
5	Wiedemann-99 CC0 Default = 1.5	0.6	0.65	<0.01	
		1.2	0.04	0.01	
		3	0.37	0.18	
6	Wiedemann-99 CC1 Default = 0.9	0.5	<0.01	<0.01	
		1	<0.01	0.01	
		1.5	0.19	0.19	
7	Wiedemann-99 CC2 Default = 4	1.5		0.06	
		3		<0.01	
		6		0.82	
8	Wiedemann-99 CC3 Default = −8	−4	0.03	0.17	0.65
		−10	0.01	<0.01	0.93
		−15	0.01	0.43	<0.01
9	Wiedemann-99 CC6 Default = 11.44	2		0.44	
		8		0.03	
		20		<0.01	
10	Wiedemann-99 CC7 Default = 0.25	0.15		0.01	0.05
		0.3		0.32	0.03
		0.5		0.49	0.18
11	Wiedemann-99 CC8 Default = 3.5	2	0.01	<0.01	
		3.6	0.87	0.88	
		5	0.69	0.75	
12	General behavior Default = Free lane selection	Right side rule	<0.01		
13	Desired position at free flow Default = middle of lane	Any right	0.02	0.10	
			0.01	<0.01	

[a]T represents the simulated total conflict count
[b]R represents the simulated rear-end conflict count
[c]L represents the simulated lane-change conflict count
[d]C represents the simulated crossing conflict count

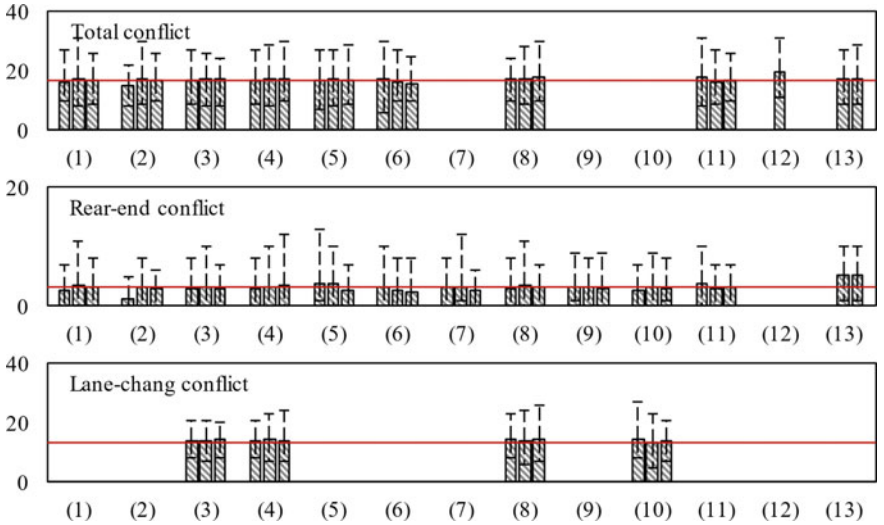


Fig. 1 Graphical descriptions of sensitive parameters for different conflict type

Fig. 1. The parameter number shown in the parentheses in Fig. 1 were the same as the parameter index in Table 2.

4.1 Following Behavior

Look ahead distance defined the distance that a vehicle obtains information, which included three parameters [11]. They were the max distance, the min distance and the observed vehicle count. As indicated by the student t-test, the max distance impacted all types of simulated traffic conflicts significantly, except for lane-change conflict. The observed vehicle count influenced the total count of simulated traffic conflict, rear-end conflict. Temporary lack of attention simply defined the situation that vehicles did not react to a preceding vehicle for a few seconds. Two parameters available in this set were the duration and the probability. Sensitivity test shown that the two parameters had impacts on all types of simulated traffic conflict significantly, i.e., a slight increase in the percentage of probability or in the time of duration resulted in a statistically significant increase in all types of simulated traffic conflict.

Another important set of parameters came from the Wiedemann's car following model in 1999. More specifically, CC0 defined the desired static distance for stopped cars; CC1 was the time that a driver wanted to keep; CC2 restricted the additional distance required by a driver in following behavior; CC3 defined the threshold for deceleration process; CC4 and CC5 constrained the allowed speed differences in following status; CC6 defined the underlying relationship between following distance and vehicle speed; CC7 was the actual acceleration during oscillation process; CC8

constrained the acceleration for standing vehicles; CC9 was the expected acceleration value at 50 mile/hr. It was found that three parameters, which were CC4, CC5, and CC9, failed to reject the null hypothesis in the Student's t-test at 95% level of confidence. More specifically, CC0, CC1, and CC8 impacted the total count of traffic conflict significantly. CC3 influenced all the three types of traffic conflict, which revealed the authors' previous study in 2013 [5]. CC7 was sensitive to both the rear-end conflict and the lane-change conflict, while CC2 and CC6 were sensitive to rear-end conflict count only.

4.2 Lang Change Behavior

The general behavior defined the two basic groups of lane selection rules in the VISSIM, which were right side rule and free lane selection. The main difference between the two groups located in the consideration of a collision time threshold. As shown in Table 2, a significantly higher total count of traffic conflict under right side rule was obtained compared to the total count of simulated traffic conflict under free lane selection at freeway merging area. However, the differences in rear-end conflicts or in lane-change conflict were not statistically significant. Moreover, the six parameters in the set of necessary lane change were used to capture the aggressiveness of lane change behavior due to fixed routing choice. As the analysis results shown, there was no significant difference posted in the conducted Student's t-tests. One potential explanation for the analysis results was that parameters of necessary lane change affected the characteristics of the conflict itself, but not the generation of simulated traffic conflict. A lower or higher value of Maximum Deceleration, Deceleration Distance or Accepted Deceleration made a difference in conflict features.

4.3 Lateral

The Desired Position at Free Flow defined the desired lateral position of a vehicle at free flow. When drivers tent to drive at right or random position of lane, they might be influenced by other drivers on the neighboring lanes, which resulted in a significant increase in the total count of traffic conflict and the count of rear-end conflict. As shown in Table 2, the Desired Position showed a significant difference in the count of simulated traffic conflict and the count of rear-end conflict.

5 Conclusions

In this study, sensitivity analysis was conducted to 31 driver behavior related parameters in VISSIM. Different parameter value was selected for each parameter according to previous studies and the practical engineering applications. Furthermore, in order to reduce stochastic variability, multiple runs were conducted for each parameter level magnitude. Student's t-tests were also conducted to the simulated conflict counts output from the SSAM. The following conclusions are based on the data analysis result:

(a) 10 parameters were sensitive to the total count of traffic conflict, and 12 parameters were sensitive to the count of rear-end conflict, and 4 parameters were sensitive to the count of lane-change conflict.
(b) For most sensitive parameters, it could be noticed that there was no significant monotonicity of the count of simulated traffic conflict when the parameters changed their value.
(c) Some parameters, which were seldom mentioned in previous studies, were also sensitive to the count of simulated traffic conflicts, including the Duration and the Probability for temporary lack of attention. A slight increase in the percentage of Probability or in the time of Duration led to a statistically significant increase in the count of all types of simulated traffic conflicts.

The study presented in this paper demonstrated the impacts of different human behavior parameters provided in the VISSIM on the simulated conflicts. The findings revealed the previous work proposing the two-stage calibration process for traffic conflicts and were in line with expectations. The results were presented to provide transportation professionals with additional insight into simulated conflict using SSAM and useful knowledge for model development and calibration. The study focused on the sensitivity of single parameters for simulated traffic conflicts only. However, further studies were needed in the sensitive analysis of multiparameters for traffic safety. In addition, Student's t-test was conducted in this study. And parameters with statistical significant differences were recognized to be sensitive. However, the impacts of sensitive parameters were not quantified. The evaluation countermeasures for the impacts of these parameters on simulated conflicts merited further investigations.

Acknowledgements The authors appreciate the funding support from the National Natural Science Foundation of China (51508122, 51478113, 51508094), Guangxi science and technology projects (1524800210, GuikeAB16380280, Guike-AB17292087), the Natural Science Foundation of Guangxi (Grant No. 2015GXNSFBA139216), the Natural Science Foundation of Jiangsu (BK20150612), Shanghai Rising-Star Program (16QB1403000), Shanghai Urban-Rural Development Transportation Talents Special Funds, as well as The Scientific research project of Chinese National Ministry of Housing and Urban-Rural Construction (2017-K2-009).

References

1. Liu P, Qu X, Yu H, Wang W, Cao B (2012) Development of a VISSIM simulation model for U-turns at unsignalized intersections. ASCE J Transp Eng 138(11):1333–1339
2. Gettman D (2003) Surrogate safety measures from traffic simulation models. FHWA-RD03050
3. Gettman D (2008) Surrogate safety assessment model and validation: final report. FHWAHRT-08-051
4. Gousios S, Garber N (2009) Relationship between time to collision conflicts and crashes on interstate highways subjected to truck lane restrictions. In: Presented at 88th annual meeting of the Transportation Research Board, Washington, D.C.
5. Fan R, Yu H, Liu P, Wang W (2013) Using VISSIM simulation model and surrogate safety assessment model for estimating field measured traffic conflicts at freeway merge area. IET Intel Transp Syst 7(7):68–77
6. Huang F, Liu P, Yu H, Wang W (2013) Identification if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections. Accid Anal Prev 50:1014–1024
7. Dijkstra A, Marchesini P (2010) Are calculated conflicts in a micro-simulation model predicting the number of crashes. Presented at 89th annual meeting of the Transportation Research Board, Washington, D.C.
8. Nicholas E, Randy B (2006) Sensitivity of simulated capacity to VISSIM driver behavior parameter modification. Transp Res Rec J Transp Res Board, 102–110
9. Chu L, Liu H, Recker W (2003) A calibration procedure for microscopic traffic simulation. Intell Transp Syst Proc 2:12–15
10. Washington S, Karlaftis M, Mannering F (2003) Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC, Boca Raton
11. PTV—Planung Transport Verkehr AG (2004) User manual VISSIM version 4.20. Karlsruhe, Germany

Research on the Analysis of Campus' Accessibility Based on Individual Activity Type

Baohong He, Xiang Zhang and Xuefeng Li

Abstract The previous accessibility model is fundamental but neglects activity type on transportation analysis and planning. It fails to comprehensively evaluate individuals' travel behaviors and their utilizable spatial-temporal resources under different activity types. This paper applies time geography method and classifies the activity types according to the characteristics and the elastic degree of individual activities. And then, the spatial temporal accessibility model and extended model based on the characteristics of different activity types are constructed. Moreover, a case study of campus trip data is given to verify the rationality of the models. The results show that when considering the type of activities, the accessibility does not follow the characteristics of scatter diagram and a "core to periphery" layer structure which is from high to low. Instead, it is determined by the type of activity. The stronger the mandatory activities are, the higher the accessibility of the region will be. Furthermore, when there are only one kind of facilities, the travel distance and time are major factors affecting the value of accessibility, and the personal selection of facilities follows the principle of proximity. The results are more consistent with the real life, therefore the proposed models are more rational. The results of this study provides great reference to the quantification of urban accessibility and theoretical support to the allocation of public urban facilities.

Keywords Spatio-temporal accessibility · Model measuring · Activity type · Spatio-temporal constraints · Personal preference

B. He (✉) · X. Zhang
Faculty of Transportation Engineering, Kunming University of Science and Technology, Yunnan Kunming 650500, China
e-mail: 94002267@qq.com

X. Zhang
e-mail: 615582070@qq.com

X. Li
School of Transportation, Southeast University, Jiangsu Nanjing 210096, China
e-mail: lixuefengseu@foxmail.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_11

1 Introduction

With the reform of Chinese higher education system and the expansion of Chinese enrollment scale, building new large-scale campus in the suburbs is becoming more popular among colleges and universities [1]. As a result of the traditional urban design, the current campus planning is more inclined to use urban design principles to create a magnificent architectural style, but ignores the space demand of students' small-scale activities [2], which leads to the longer walk between school functional areas, longer traffic flow, fewer social communicative activities and other issues which have seriously affected the students' learning and living quality [3]. Accessibility is a key indicator which reflects the coordinated relationships among transportation system, land use, and other urban infrastructures [4, 5]. Also, accessibility can be used to reveal the interrelationship between land use and individual's demand, and helps further the understanding of the interrelation between transportation planning and travel behavior, which is an important tool to evaluate the reasonability of the urban planning [6].

In the past, because the relationship of people, places, and activities were more limited by distance, the key location (such as households or workplaces) could be used to infer people's activity characteristics. However, the time-space compression makes the relationship among people, land and activities more complex. People can do more activities in one or more places, or do the same activities in different places. This new mode of communication changes individuals' characteristics and activities [7]. With the development of GIS technology [8], spatio-temporal resources are considered when accessibility is measured [9–11]. However, in real life, the activities of individuals are not only limited by time and space, but also influenced by the types of individuals' activities. Some activities (such as work, school) must be completed at a fixed time and place. Some activities (such as shopping, dining, bus) can be completed at a certain time or place, and others (such as entertainment, sport) are not. Current accessibility measuring model cannot illustrate the complex behavior of individuals due to different types of activities. Also, current models lack the consideration of their activity preferences which was decided by people themselves.

This paper first classifies the activity types according to the activities characteristics and the elastic degree. Activities that have less limitation with individuals' desire and demand, and completed at a confirmed time and place are defined as mandatory activities. Activities affected by individuals' desire, and the time and space is alternative to some extend are defined as optionality activities. Activities that are not limited to the time and space restriction and can be completed in accordance with individuals' desires are defined as unconstrained activities. This paper creates a different spatio-temporal accessibility measuring models and extension model under the characteristics of these three activities. The modes can analyze the relationship between urban residents' accessibility and their travel behavior under different types of activities, and provide reference for urban traffic planning and design. Also, it provide a new idea for transport policy makers.

2 Limitations of Previous Accessibility Measures

2.1 Accessibility Measure Base on Location

Hansen proposed the concept of accessibility and defined it from the point of view of interaction chance [12]. Afterwards, many scholars proposed a variety of spatial models based on the concept of accessibility and carried out a large number of empirical research. Ingram proposed the concept of relative accessibility and described it by spatial barrier that consist of distance, time and cost, creating spatial barrier model of accessibility [13]. Baxter, Lenisand Kirby modified the spatial barrier model. Since there is a lack of the consideration of the chance to approach the starting point [14, 15], Church et al. proposed a maximum coverage model based on the minimum distance model, which measures the accessibility level by assessing the local of travel [16]. Because the maximum coverage model and the spatial barrier model do not consider the interaction between the measurement point and the attraction point as well as the attenuation of the spatial effect on distance, William proposed the gravitational model of the public service facility by reference to the concept of gravitational model in physics. By calculating the "gravity" between the public service facilities and the demanders [17]. Although the gravitation model has been widely used, there are still some factors not considered such as population distribution, residents' demand and choice. Joseph added an impact factor of population-scale to the gravitational model to reflect the effect of population size differences at different settlements on the facility layout efficiency [18]. Rafael used a gravitational model to study the location of supermarkets in the Gran Canaria islands by adding a proportional selection rule to reflect facility service capacity and service quality [19]. In order to overcome the problem that the results of gravity model cannot calculate the results, which is incomparable due to the inconsistency of the number of attraction points [20], the probabilistic method and frequency method of accessibility are developed based on the gravity model to modify the force of gravity model [21]. Peeters summarized the results of Love et al, and concluded the coefficient of travel friction within [1.5, 2] has little impact on the results after many researches [22]. Based on these modes, the constrained gravitational model was proposed. Wilson used the maximum entropy law of statistic to measure accessibility, including the single constraint gravitational model and the double constrained gravitation model [23, 24]. Furthermore, Ben-Akiva M and Lerman SR proposed a accessibility model based on stochastic utility, which considers accessibility relating to personal trip choice [5]. Schürman used grid technology to describe the spatial characteristics of the daily potential index of railways, highways, and airways. It has the advantage of delineating the service scope of traffic stations and can better show the accessibility details of regional spatial patterns [25, 26]. Since 1990s, GIS technology is extensively used, Langford and colleagues used a two-step mobile search method to study the impact of population distribution on accessibility in the Cardiff region with GIS software [27].

All these accessibility measures emphasize the geospatial spatial characteristics. In the accessibility model, there is a lack of consideration in space-time factors and

individual factors. This led to the lack of consideration for people who are active in the geographic space. Also, it regarded the individual demand match for facilities and the time of using facilities unrestricted. It cannot scientifically reflect the impact of human space-time characteristics on accessibility.

2.2 Time and Space Accessibility Measures

Space-time accessibility measures are based on the construct of the space-time prism proposed by Hägerstrand [28] and elaborated by Lenntorp [29], Space–time accessibility measures enhance the previous approaches by reducing the set of feasible activity locations with the spatiotemporal constraints of individuals. Researchers also begun to improve the model of temporal and spatial accessibility measurement, except for the limit caused by space and time resources. Miller concluded that the individual's temporal and spatial accessibility are also influenced by the duration and travel time of activities [30], Burns [29], Hsu and Hsieh [31], Chen and Kwan [32] also emphasized the need to consider individual trip chains and activity patterns when measuring temporal and spatial accessibility. Based on these research, the time and space accessibility measure model has been further optimized. Ennio Cascetta and Armando Cartenìre provide a new behavioral definition of accessibility, rooted in an opportunity-based framework, and a corresponding accessibility model that expresses accessibility as the number of "available" opportunities, behavioral attributes are taken into framework [33]. Hyun-Mi Kim, Mei-Po Kwan seeks to enhance space-time accessibility measures through developing a new operational method and GIS-based algorithm that better represents the space-time characteristics of urban opportunities and human activity-travel behavior [8]. Fang Ren, Daoqin Tong, Mei-Po Kwan analyzed recent developments in space–time measures of individual accessibility to explore the spatial and temporal structures of demand by considering individuals' space–time constraints and impact of existing urban structures [34]. In these approaches, attractiveness of activity locations is a function of time-related attributes alone. Also activity attributes are not taken into account. This is a major limitation of the time-space accessibility measures in current models.

In real life, some activities are required to be done at a fixed time and place (such as work, school); some activities need to be completed either at a fixed time or place (shopping or dining), and others are not subject to time and space resources. In addition, the current time-space accessibility measurement model cannot explain the complex behavior which influenced by the activity type of the individual. Different types of activities are important factors that lead to the change of individual temporal and spatial accessibility, which will lead to different travel characteristics of activities. Therefore, exploring the relationship between individual spatial-temporal accessibility and behavioral characteristics under different types of activities are of great importance.

3 The Elastic Characteristics and Types of Students' Daily Activities

3.1 Analysis of the Elastic Characteristics of Students' Daily Activities

Because of the special function planning of the university, it is undoubted that learning is the center of the student' main activities and the campus activities are based on the curriculum. The main body of school life is learning and include dining, shopping, entertainment, sports, social activities and others. Based on Maslow's Hierarchy of needs, this paper proposed the concept of the degree of elasticity of activity, which refers to the freedom degree of an individual practice in an activity under the constraints of space and time resources. The degree of elasticity of activity reflects the space constraints and opportunities of individuals. It determines mobility and accessibility of time and space to some extent. Therefore, this paper combined the spatial characteristics of students' daily activities, and distinguished the activity elasticity grades by the time and space constraints. The elasticity's of the campus activities are shown in Table 1.

Strong constraints: the temporal dimension and spatial dimension of activity are strongly restricted by time and space, that is to say the activity start and ending time and duration cannot be corresponding to their own aspiration. Moderate constraints generally indicate that either the temporal dimension or spatial dimension of activity is strongly limited by temporal and spatial resources, in this case, the time and duration of the activity can be flexibly arranged according to their own decisions, and the activity space is alternative to some extent. Weak constraints: the time dimension or space dimension are not strongly constrained by time and space and the duration of activities, starting and ending time and the choice of activity destination can be completed in accordance to their own demand or wishes.

3.2 Type Division of Students' Daily Activity

By analyzing the elasticity of students' daily activities, this paper used activity elastic degree as the index to judge activity types. According to the degree of activity flexibility and spatial and temporal characteristics, the campus daily activities can be divided as mandatory activities, optionality activities and unconstrained activities. Mandatory activities refers to the strong constraints of time and space, whose activity flexibility is low. It is generally not influenced by the subjective wishes and has the characteristic of a clear anchor. Namely, the activities need to be completed at the fixed time and place (such as class, listening to lectures). For the optionality activities the flexible activity is not totally limited by time and space resources. The time dimension of flexible activity is influenced by the opening time of facilities.

Table 1 Statistical tables of elastic intensity on campus activities

Activity name	Class	Lectures	Meals	Express delivery	Shopping	Taking bus	Sports	Entertainment and social interaction
Time dimension	Fixed	Fixed	Certain restrictions	Certain restrictions	Certain restrictions	Certain restrictions	Freedom	Freedom
Space dimension	Fixed	Fixed	Freedom	Freedom	Freedom	Freedom	Freedom	Freedom
Elasticity of activity	Strong constraint	Strong constraint	Moderate constraints	Moderate constraints	Moderate constraints	Moderate constraints	Week constraint	Week constraint

Table 2 The characteristics of different types of activities

Activity type	Degree of elasticity	Activity time dimension	Activity place dimension	Nature of activity	Activity
Restricted activities	Strong constraint	Activity start and end time	Single event location	High density, universality	Class, lecture
Resilient activity	Moderate constraint	Opening hours	With a few substitutability	A certain randomness	Meals, express shopping, taking bus
Free activity	Week constraint	Almost a whole day	With a bit strong substitutability	A bit strong randomness	Sports, entertainment and social interaction

Meanwhile, the location of facilities can be replaced. Unconstrained activities has the least restriction to space-time resources. The opening time of facilities has little effect on the temporal dimension of activity flexibility. Meanwhile, the facilities can be substituted for higher places and have stronger freedom. It can be completed according to individuals' subjective wishes (Such as social, sports, leisure, entertainment), with strong arbitrariness and uncertainty. The types and characteristics of different activities are shown in Table 2.

4 New Temporal-Spatial Accessibility Model of Individual Activity Type

4.1 The Time and Space Accessibility Model for Mandatory Activities

Mandatory activity is an activity that must be completed by the individual in the activity plan of the day. It cannot be influenced by the location and the time of the activities. At the same time, it is necessary to complete the activity at the appointed time and place which is irreplaceable. Meanwhile, the start and the end time of the activity and the minimum duration of the activity are fixed. Besides, it is essential to take into account that the travel willingness will decrease with the increase of travel time. The constraint model based on the gravitational model can be expressed as follows:

$$A_{ik} = e^{-\lambda t_{ik}} \cdot T(POI_i) \tag{1}$$

where: A_{ik} is the accessibility measure of the mandatory activity site K relative to the individual i;t_{ik} is the travel time; λ is the distance attenuation coefficient, which indicates the distance attenuation sensitivity between the origin and destination; $e^{-\lambda t_{ik}}$ is the negative exponential distance decay function. This function can better simulate the individual travel intention, reflecting the individual i in the actual travel process whose travel desire will weaken with the increase in travel time. $T(POI_i)$ is the activity time of the individual i at the active place.

4.2 The Time and Space Accessibility Model for Optionality Activities

Optionality activity is an activity done by the individuals in accordance with their own preferences and demand in a day to some extent. The willingness to travel will be weakened with the increase of time, and the place of activity is replaceable. Because the activity facilities may be replaced by other facilities with similar properties, it is not only need to consider the attractiveness of the facility, but also consider whether the individual's availability of the time is consistent with the availability of the facility. Therefore, Compared with mandatory activity, optionality activity has more constraint conditions when the individual time and space accessibility model is used to express the accessibility of such activities. The individual time and space accessibility model of optionality activity can be shown as follows:

$$B_{ik} = D_{ik} \cdot a_{ik} \cdot e^{-\lambda t_{ik}} \cdot T(POI_i) \tag{2}$$

Among them,

$$D_{ik} = \begin{cases} 1, T(POI_i) \leq t_1 - t_2 \\ 0, T(POI_i) < t_1 - t_2 \end{cases}$$

$$a_{ik} = w_k \cdot P_{ik}$$

$$P_{ik} = \begin{cases} 1, type_k \in Pre_i \\ 0, type_k \notin Pre_i \end{cases}$$

where: B_{ik} is the accessibility measure of the optionality activity site K relative to the individual i; D_{ik} is the judgment condition of the time and space constraints, and is used to judge whether the utilization time interval of the individual i at the active site K is within the time when the facility is open; when the time use interval of the individual i in the active location K among the open time takes 1, otherwise 0; a_{ik} is the attraction factor of activity point K, indicating the degree of attraction of activity point K to individual i, and w_k is the attraction coefficient of active site. P_{ik} is a judgment condition for personal preference, determining whether the type $type_k$ of the activity point K belongs to the activity type set Pre_i of the individual preference,

and Pre_i is set by the individual. t_{ik} is the travel time; λ is the distance attenuation coefficient, and $T(POI_i)$ is the activity time of the individual i at the active site.

4.3 The Time and Space Accessibility Model of Unconstrained Activities

Unconstrained activities are rarely affected by the opening hours of the facilities, and individuals are less constrained by space and time. There is a strong alternative to the location of the activities, and the individuals tend to choose the public facilities with fully function to participate in such activities Meanwhile, travel desire will be reduced with the increase of travel time and distance. Therefore, when individual time and space accessibility is used to represent this type of accessibility of individual unconstrained activities, it needs to consider the attraction of the facilities, travel attenuation, individual preferences and the individual in the active location of the use of time. The individual time and space accessibility model of unconstrained activities can be shown as follows:

$$\begin{aligned} C_{ik} &= a_{ik} \cdot e^{-\lambda t_{ik}} \cdot T(POI_i) \\ a_{ik} &= w_k \cdot P_{ik} \\ P_{ik} &= \begin{cases} 1, type_k \in Pre_i \\ 0, type_k \notin Pre_i \end{cases} \end{aligned} \tag{3}$$

where, C_{ik} is the accessibility measure of the unconstrained activity site K relative to the individual i; a_{ik} is the attraction factor of activity point K, indicating the degree of attraction of activity point K to individual i and w_k is the attraction coefficient of active site. P_{ik} is a judgment condition for personal preference, determining whether the type $type_k$ of the activity point K belongs to the activity type set Pre_i of the individual preference, and Pre_i is set by the individual. t_{ik} is the travel time; λ is the distance attenuation coefficient, and $T(POI_i)$ is the activity time of the individual i at the active site.

4.4 Extension of the Model

Based on the measurement of time and space accessibility on the above-mentioned different activity types, this paper used A_i^{max} to represent the preference activity of an individual to assess the individual's living condition. Using the cumulative measure of activity A_i^{accu} to measure the degree of freedom of guiding and carrying out various activities by free personal preference when time and space resources of the individual or group are limited. And then, it serves the assessment of individual or

group social life status and social group time and space accessibility. The expressions for A_i^{max} and A_i^{accu} are shown in Eqs. (4) and (5):

$$A_i^{max} = \max(A_{ik}, B_{ik}, C_{ik}) \tag{4}$$

$$A_i^{accu} = A_{ik} + B_{ik} + C_{ik} \tag{5}$$

5 Case Study

5.1 Research Data

This paper takes the new campus of Kunming University of science and technology in Cheng Gong as an example to select three dormitory areas located in different region for field investigation. The campus is divided into three groups: Yi Yuan, Jing Yuan and Tian Yuan. The library and teaching area are all located in the west of the avenue, and the three dormitory areas and logistics centers are in the other side. There are two main ways to obtain data in this research: one is to get the relevant information of campus planning layout, and public facilities through the school website, reference room, the other is to conduct questionnaires and surveys in school in order to investigate and analyze the daily life of the students in the three dormitory areas of the campus. The survey contained a total of 203 valid samples. In these samples, men are slightly more than women. Meanwhile, the percentage of Yi Yuan, Jing Yuan and Tian Yuan in samples are 33%, 30% and 37% respectively. The content of the questionnaire included gender, grade, living dormitory areas, transportation mode, activity-based travel time, duration of activities, travel frequency and other information. The questionnaire data were roundly descriptive analyzed by statistical analysis software, which was used as the basis of the research. Campus space structure as shown in Fig. 1.

Judging from the sample statistics of data, the men accounted for 57.65% of the sample, and the percentage of women is 42.36%. It is obvious that men are slightly more than women. Which is in line with the phenomenon of men more than women in China's engineering school. From the daily travel mode of the students, bicycle and motor vehicle travel is accounted for 19.66% of all the male in ways of traveling, and bicycle or motor vehicle travel is accounted for 8.14% of all the female in ways of traveling. The proportion of the female campus bus traveling is greater than the male, especially in the middle and long distance travelling. On the contrary, the demand for bicycles and electric vehicles is reduced. The travel frequency of students from Yiyuan dormitory area is higher than other dormitory areas in daily trip. Due to the perfect public facilities. The group in Yiyuan living quarters take more active part in the outings which are optionality activities and unconstrained activities. Travel frequency of each living area is shown in Fig. 2.

Fig. 1 Campus space structure

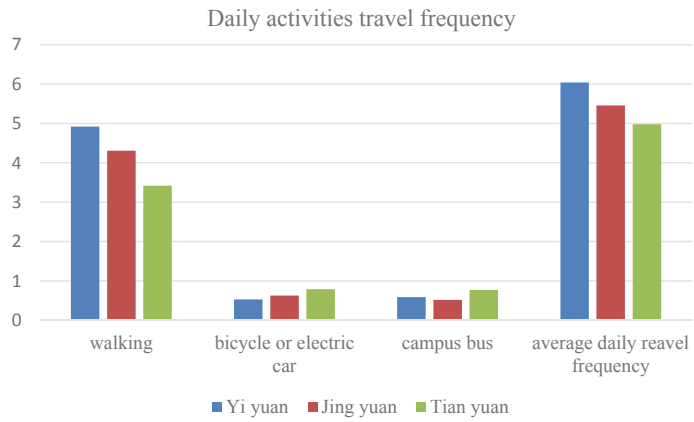


Fig. 2 Travel frequency of each living area

5.2 Model Parameters

In the traditional accessibility measuring model, there is not a uniform standard for the attraction coefficient of the facilities, and it is usually replaced with the parameters such as "facility scale", "facility quality", "facility grade" [35]. However, it is difficult to measure the scale of the public facilities with objectively in real life. On the one hand, the usage frequency of the activity sites reflects the students' actual usage rate of campus public facilities. On the other hand, it shows the demand degree of students for facilities with different quality. So this paper used the usage frequency of the facilities to express attraction coefficient w_k of the public facilities, the usage frequency of the classroom is not included in the statistics, because it is not affected by the student's subjective initiative as the classroom is the place for the class. Through statistical analysis, the average frequency of students' usage of public facilities on campus is shown in Table 3.

The attractiveness of public facilities is indicated by the average usage frequency of public facilities on campus. When the frequency is more than 1, $w_k = 0.3$. When the frequency is within 1 to 10, $w_k = 0.2$. When the frequency is less than 1, $w_k = 0.1$. Distance decay function describes the travel intention of individual will gradually decay with the increase of travel time or travel distance during the travel process, especially in the low cost range. While the decay is relatively smooth among the high cost range. It is an important parameter to study the space-time accessibility of individuals. The value of γ in the model has been studied by foreign researchers is $\gamma = 0.9$. In the study of the space-time accessibility change of urban bus trip mode in Guangzhou by scholar Hu Jihua, Li Guoyuan and Zhong Guangpeng [36], value of the exponential function is $\gamma = 0.9$. Therefore, the power function of the distance decay function of the gamma value is $\gamma = 0.9$ in this paper after studying the relevant literature.

Table 3 Average frequency of public facilities on campus

The name of public facilities	Canteen	Supermarket	Express station	Playing field	Library	Self-study room	Bus stop
Average frequency (times per week)	22	7.6	0.8	2.2	2.5	0.6	2.1

6 Spatial Distribution Characteristics of Spatial Accessibility

In this paper, we treat the facilities with unified characteristic as a particle. The living area of the dormitory is treated as the center, and the corresponding parameters (attraction coefficient of the campus public facilities, distance decay function,travel time, available time of the individual in the interest points, starting time and the closing time of the facilities, etc.) of the model are introduced into the formula for spatial analysis, Based on it, The combined formula (1–5) can be calculated in ArcGIS9.0. The time and space accessibility distribution of campus is shown in Fig. 3. The results are consistent with the present situation of campus, which verified the rationality of the model.

The larger the bubble in Fig. 2, the higher the accessibility. There is a common feature in the three living areas, when considering the type of activity. When considering the type of activities, the accessibility isn't showing the characteristics of scatter diagram and a "core to periphery" layer structure which is from high to low. Also, there is no appearance that the more close to the living areas, the higher the accessibility. But for teaching areas, it is the highest in three living areas, followed by the library. The accessibility level of two learning region is significantly higher than other regions. Not only because the facilities are complete, but also due to its high degree of demand, the individual preference in these places is stronger than other areas, so as to complete the class, self-study and other activities, students can overcome the external resistance, thereby increasing accessibility. Because of the less attraction to students, the accessibility in these area such as playing field is relatively weak. The phenomenon above is more consistent with the life mode of the students which learning is the main line. When compared with other living areas, the results

Fig. 3 Distribution of time-space accessibility from different living space

show that it is obvious were significant differences in the accessibility of the same facilities for different living areas. The accessibility of the teaching area, playing field and other facilities in Yiyuan is higher than correspondence facilities in other two living areas. It indicates that the distance and time of the same facilities have essential influence on the accessibility. In addition, the accessibility of the living facilities such as supermarket, canteen, bus stops, and express post in every area is relatively higher than that of other facilities in the region. It shows that the mandatory activities with the highest degree of demand are less constrained by the external environment. There will be a phenomenon of cross regional completion, even by overcoming the external resistance (such as changing the mobility of traffic); For optionality activities and unconstrained activities, the individual will generally follow the principle of proximity in the distance from their nearest facilities to practice activities. However, More complete facilities in the area will also induce more optionality activities and unconstrained activities, but will not induce the mandatory activities.

6.1 Suggestion of the Campus Space Planning

Accessibility is affected by many factors through the above analysis, but the activity type is an important factor that cannot be ignored. Therefore, this paper makes suggestions on campus planning from the perspective of individual activities, it reflects the relationship between the unit of activity and the activity space, considering the dynamic relationship between the demand of individuals and supply level of public facilities, so as to ensure the fairness of the public facilities space planning and social rationality. Based on the concept of life circle, campus space planning should combine the characteristics of students' activities and satisfy the basic demand of their life firstly, such as clothing, food, shelter and travel. Next, the planning should meet the demand of activities related to learning. Finally it seeks to meet the spiritual demand of students, including leisure, sports, social activities or other higher demand. In this paper, the daily life circle of the campus is divided into three types: basic living circle, learning life circle and entertainment social life circle. The campus daily life circle planning is aimed at demand behavior of students and mainly to organize campus living space, public facilities and reasonable allocation of campus resources, which has a great significance for the implementation of people-oriented campus planning. Campus daily life circle shown in Fig. 4.

The traditional campus planning of China is affected by the idea of function division, which focus on the large scale space planning and facilities supply, from the perspective of individual activity types and demand characteristics, based on the transformation from the space dominated by teaching to the spatial pattern of student demand, Based on the individual demand of the life cycle, the campus space planning is shown in Fig. 5.

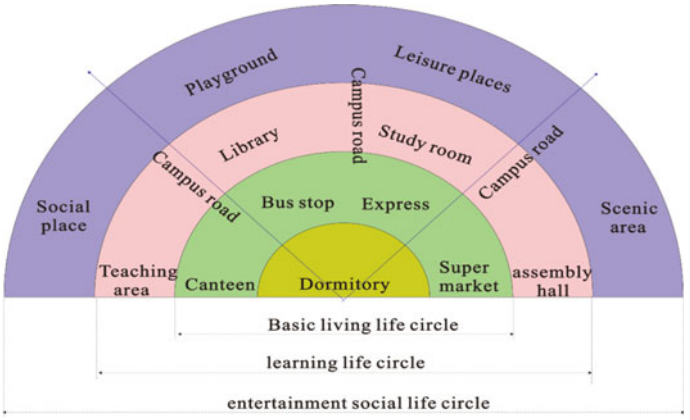


Fig. 4 Sketch map of the campus life circle

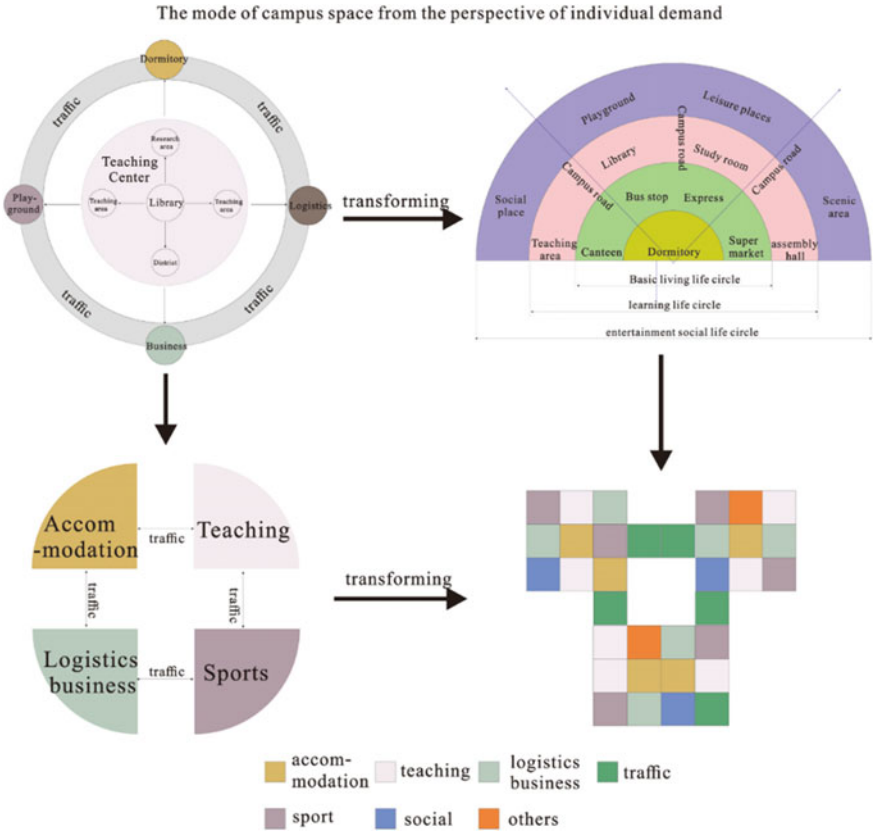


Fig. 5 Campus space from the perspective of individual demand

7 Conclusions

The previous accessibility model often neglected the influence of the activity types. Individual activities are constrained by time and space resources. Also these activities will change the travel behavior that are influenced by the type of activity. This is an important factor that cannot be ignored when calculating accessibility model. In order to make the calculation results more consistent with the real life situations, the activity type is divided into three categories according to activities characteristics and the degree of elasticity. Namely: mandatory activities, optionality activities and unconstrained activities. According to the characteristics of each activity and the individual's preference, this paper proposed a model of time and space accessibility, and took a campus as a case study. Results showed that when considering the type of activity, the distribution characteristics of accessibility does not follow the scatter diagram and the layer structure is not the circle whose" core to periphery" is from high to low. On the contrary, the stronger the mandatory activities is, the higher the accessibility of the region will be. However, the accessibility of the optionality activities and unconstrained activities of region is relatively weak. This is due to the promotion of accessibility with the overcoming of external resistance when people participate in mandatory activities. Therefore, accessibility is the carrier of the transformation between the three resources of time and space. Also, it is the individual's own factors that can improve the level of accessibility. Moreover, there is great differences among the accessibility of the same facilities for different living areas, which is influenced by the distance or time of travel. To meet their demand, individuals' selections of facilities follow the principle of proximity, especially for the optionality activities and unconstrained activities. Therefore, cross-regional travel is common in the mandatory activities but rare in optionality activities and unconstrained activities. Finally, according to the students' individual demand, the paper proposed a campus planning model which is based on the diversity of life circle.

The innovation of this paper is that it proposed the accessibility model considering the activity type of individual. This method takes the constraints of individual and time resources into account, and can reflect the actual situation. However, the activity of individual in a day is continuous and dynamic, which is in the performance of travel chain. The starting point of each activity is not always the place of residence, so the OD point is dynamic. In the future work, we will describe the whole process of activity from the view of trip chain to determine the origin-destination of each activity, so that the result of the accessibility is more accurate. Moreover, we still need to consider the dynamic change of the accessibility of individuals under the action of multi person cooperation, which makes the calculation results more accurate, so as to provide reference for urban planning and transportation planning.

Acknowledgements Supported by the National Natural Science Foundation of China (51668029).

References

1. Lizhang QHC (2010) Study of temporal and spatial characteristics of students behavior in Hangzhou Xiasha higher education eastern park. Geogr Res 29(7):1281–1290
2. Chen Z (2014) From production space to living space: change of urban function and corresponding spatial planning strategy. City Plan Rev 38(4):28–33
3. Shi LGW (2014) Analysis on the traffic optimization strategies of large-scale campus: a case study of Chenggong Campus, Kunming University of Science and Technology. Huazhong Archit 9:131–135
4. Odoki JB, Kerali HR, Santorini F (2001) An integrated model for quantifying accessibility-benefits in developing countries. Transp Res Part A 35(7):601–623
5. Ben-Akiva ME, Lerman SR (1979) Disaggregate travel and mobility choice models and measures of accessibility
6. Wang Y, Monzon A, Ciommo FD (2014) Assessing the accessibility impact of transport policy by a land-use and transport interaction model—the case of Madrid. Comput Environ Urban Syst 49:126–135
7. Miller H (2007) Place-based versus people-based geographic information science. Geogr Compass 10(3):503–5351749
8. Kim HM, Kwan MP (2003) Space-time accessibility measures: a geocomputational algorithm with a focus on the feasible opportunity set and possible activity duration. J Geogr Syst 5(1):71–91
9. Neutens T, Delafontaine M, Schwanen T et al (2011) The relationship between opening hours and accessibility of public service delivery. J Transp Geogr 25(3):128–140
10. Neutens T, Delafontaine M, Scott DM et al (2012) A GIS-based method to identify spatiotemporal gaps in public service delivery. Appl Geogr 32(2):253–264
11. Neutens T, Schwanen T, Witlox F et al (2009) Equity of urban service delivery: a comparison of different accessibility measures. In: Proceedings of the nectar cluster 6 meeting on accessibility, policy making, spatial planning
12. Hansen WG (1959) How accessibility shapes land use. J Am Planning Assoc 25(2):73–76
13. Ingram DR (1971) The concept of accessibility: a search for an operational form. Reg Stud 5(2):101–107
14. Baxter RS, Lenzi G (1975) The measurement of relative accessibility. Reg Stud 9(1):15–26
15. Kirby HR (1976) Accessibility indices for abstract road networks. Reg Stud 10(10):479–482
16. Church R, Revelle C (1974) The maximal covering location problem. Pap Reg Sci 32(1):101–118
17. Shonick W (1976) Elements of planning for area-wide personal health services. Mosby
18. Joseph AE, Bantock PR (1987) Measuring potential physical accessibility to general practitioners in rural areas: a method and case study. Williams & Wilkins and Associates Pty
19. Suárez-Vega R, Santos-Peñate DR, Dorta-González P et al (2011) A multi-criteria GIS based procedure to solve a network competitive location problem. Appl Geogr 31(1):282–291
20. Geertman SCM, Eck JRRV (1995) GIS and models of accessibility potential: an application in planning. Int J Geogr Inf Sci 9(1):67–80
21. Sakkas N, PéREZ J (2006) Elaborating metrics for the accessibility of buildings. Comput Environ Urban Syst 30(5):661–685
22. Peeters D, Thomas I (2000) Distance predicting functions and applied location-allocation models. J Geogr Syst 2(2):167–184
23. Wilson AG, Wilson AG (1971) A family of spatial interaction models, and associated developments. Environ Plan A 3(1):1–32
24. Wilson AG (1967) A statistical theory of spatial distribution models. Transp Res 1(3):253–269
25. Spiekermann K, Wegener M (1996) Trans-European networks and unequal accessibility in Europe. 4:35–42
26. Vickerman R, Spiekermann K, Wegener M (1999) Accessibility and economic development in Europe. Reg Stud 33(1):1–15

27. Langford M, Higgs G, Radcliffe J et al (2008) Urban population distribution models and service accessibility estimation. Comput Environ Urban Syst 32(1):66–80
28. HäGERSTRAND T (1970) What about people in regional science? Pap Reg Sci 24(1):6–21
29. Lenntorp B (1976) Paths in space-time environments: a time-geographic study of movement possibilities of individuals. Lund Stud Geogr 44
30. Miller HJ (1998) Measuring space-time accessibility benefits within transportation networks: basic theory and computational procedures. Geogr Anal 31(1):1–26
31. Hsu CI, Hsieh YP (2004) Travel and activity choices based on an individual accessibility model. Pap Reg Sci 83(2):387–406
32. Chen X, Kwan MP (2012) Choice set formation with multiple flexible activities under space-time constraints. Int J Geogr Inf Sci 26(5):1–21
33. Cascetta E, Cartenì A, Montanino M (2016) A behavioral model of accessibility based on the number of available opportunities. J Transp Geogr 51:45–58
34. Ren F, Tong D, Kwan MP (2014) Space-time measures of demand for service: bridging location modelling and accessibility studies through a time-geographic framework. Geogr Ann Ser B Hum Geogr 96(4):329–344
35. Sun Y, Lv B, Zhao Y (2015) A study of county public service facilities distribution assessment based on behavior investigation and GIS: a case study of medical facilities in Dexing. Hum Geogr 3:103–10
36. Hu J, Li G, Zhong G-P (2014) Measuring space-time accessibility within bus network based on space-time process of bus. J Transp Syst Eng Inf Technol 14(4):146–153

Dynamic Programming Approaches for Solving Shortest Path Problem in Transportation: Comparison and Application

Xuan Li, Xiaofei Ye and Lili Lu

Abstract This paper seeks to investigate the performance of two different dynamic programming approaches for shortest path problem of transportation road network in different context, including the Bellman's dynamic programming approach and the Dijkstra's algorithm. The procedures to implement the two algorithms are discussed in detail in this study. The application of the Bellman's approach shows that it is computationally expensive due to a lot of repetitive calculations. In comparison, the Dijkstra's algorithm can effectively improve the computational efficiency of the backward dynamic programming approach. According to whether the shortest path from the node to the original node has been found, the Dijkstra's algorithm marked the node with permanent label and temporal label. In each step, it simultaneously updates both the permanent label and temporal label to avoid the repetitive calculations in the backward dynamic programming approach. In addition, we also presented an algorithm using dynamic programming theory to solve the K shortest path problem. The K shortest path algorithm is particular useful to find the possible paths for travelers in real-world. The computational performance of the three approaches in large network is explored. This study will be useful for transportation engineers to choose the approaches to solve the shortest path problem for different needs.

1 Introduction

To find the shortest paths between different locations on a transportation network is a fundamental problem for traffic engineers to analyze the network flow [25]. Graphically, the shortest path problem is to find a path between two vertices (or nodes) in a graph such that the sum of the weights of the used edge is minimum among all feasible paths. Note the transportation network in real world is large

X. Li (✉) · X. Ye · L. Lu
Faculty of Maritime and Transportation, Ningbo University, Ningbo, China
e-mail: lixuan@nbu.edu.cn

National Traffic Management Engineering & Technology Research Centre Ningbo University Sub-centre, Ningbo, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_12

which usually involves hundreds or even thousands of nodes and links, and for most applications (e.g., traffic assignment), the algorithm for shortest path problem needs to be implemented repetitively, thus, the computational efficiency of the shortest path algorithm is very critical. In history, various algorithms have been proposed to solve the shortest path problem effectively and efficiently, including the Bellman-Ford algorithm [3], the Dijkstra's algorithm [7], the A* search algorithm [14], the Floyd-Warshall algorithm [10, 22]. Among these algorithms, the A* search algorithm is a heuristic algorithm that seeks to solve the shortest path problem for single pair of origin and destination (OD). The Floyd-Warshall algorithm seeks to find the shortest path for all OD pairs based on adjacency matrix. The Bellman-Ford algorithm and Dijkstra's algorithm are dynamic-programming-based approaches that seeks to find the shortest paths from one node to all other nodes. While those methods can solve the shortest path problem in different context effectively, new better approaches are still of keen interest by researchers. Recent studies like [18] develops an algorithm that significantly improve the computational effort to find the shortest paths between all OD pairs compared with the Floyd-Warshall algorithms. This algorithm is developed upon the component hierarchy approach developed for integer-weighted shortest path problem for undirected graphs [21]. Goldberg [11] improves the multi-level bucket shortest path algorithm proposed by Denardo and Fox [6] whose computational time is a linear function of number of nodes in the network. Other shortest paths can be seen by as Seidel [20], Feillet et al. [9], Pettie [18].

Similar to the shortest path problem, the K shortest path problem also attract a lot of attentions. The K shortest path problem is to find a set of paths between a given pair of nodes whose travel cost are smaller compared with other paths not in this set. In literature, many algorithms have been proposed in the last 40 years, (see e.g. Yen [24]; Eppstein [8]; Aljazzar and Leue [1]). Eppstein [8] develop a new algorithm by taking advantage of dynamic programming. It reduces the computational time significantly to find the K shortest paths from a given source to each vertex in the graph. Byers and Waterman [4] also developed a dynamic-programming- based algorithm to solve the K shortest path problem. This method can solve the K-shortest path problem efficiently in a directed acyclic network. In Sect. , we will introduce the method in detail.

Due to the advances of geographic information systems (GIS) technology, transportation network analyses in a GIS environment has been applied to many areas. Now the researches are more interested in developing algorithms that can be imbedded into the commercial software. To do so, these algorithms must be capable to deal with the time dependent shortest path problem as the travel time in the real network changed dynamically instead of fixed. Also in practices, more practical concerns are considered such as the waiting time in intersection, the U turns, speed limit et al. [5]. The latest research in this filed can be founded by Bauer and Delling [2], Schulz et al. [19], Holzer et al. [15], Maue et al. [17], Güner et al. [12, 13].

This paper explores 3 dynamic programming approaches for the shortest path problem. The first dynamic programming is Bellman's algorithm. The application of this algorithm finds that in every iteration, it involves many repetitive calculations which are already done before. The Dijkstra's algorithm occupies the drawback

effectively by assigning each node with a temporal and a permanent label. In each iteration, the Dijkstra's algorithm just updates the shortest path of the nodes with temporal label that is connected with newly added node with permanent label. The third dynamic programming approach is proposed by Byers and Waterman [4] to solve the K shortest path problem. Applicability of the method in large network is explored.

The structure of the paper is as follows: following the introduction part, the Bellman's dynamic programming approach is presented and the drawback of the approach is explored with a numerical example. Then the Dijkstra's algorithm is presented in detail. The computational efficiency of Bellman's approach and Dijkstra's algorithm in solving the large network shortest path problem are compared. Section 3 presents the dynamic programming approach for solving K shortest path problem. The last part concludes the paper.

2 Dynamic Programming Approach for Shortest Path Problem

2.1 Brief Introduction of Bellman's Dynamic Programming

Dynamic programming is a very powerful algorithmic paradigm where a complex problem is solved by sequentially solving a collection of subproblems. It uses the answers to small problems to figure out the larger ones, until the whole lot of them is solved. The main idea of dynamic programming is based on principle of optimality which states that regardless of the initial state and the initial decision, the remaining decision of an optimal policy must contain another optimal policy starting from the intermediate state. To treat a problem by dynamic programing, the problem should satisfy the following three requirements:

(1) the state of the system must be describable by a small number of variables.
(2) the function of a decision must be assigned to these system with different numerical values.
(3) the history of the system must have no influence on future behavior.

2.2 Bellman's Dynamic Programing for Shortest Path Problem

One main reason for the popularity of dynamic programing is that it is one of the most important and useful algorithms available for generating (exact) optimal solutions to a large class of shortest path problems. As denoted in the introduction part, there are a lot of methods developed for solving shortest path problem by using or partly

using the idea of dynamic programming. The following is a brief introduction of the procedures for implementing the Bellman's dynamic programming approach to solve the shortest path problem.

For a given road network (N, A) where N denotes the number of nodes in the system, and A denotes the number of links, the one source shortest path problem is to find the shortest path from one specific vertices to all other vertices. Assume the start vertices is u, and $f(i)$ denotes the shortest distance from the original vertices N to vertices i, then we have $f(u) = 0$. We use J to denote the set of vertices whose shortest path to the original vertices are found, and use I to denote the set of vertices whose shortest path have not been found, i.e. $I = \{N \not\subset J\}$. Note in the beginning, we have $J = \{u\}$, since the shortest path from vertices u to vertices u has been founded.

Start from the vertices in set J, the dynamic programing seeks to find the vertices in set I it should go to such that the distance between them is the shortest. Then we have the following transition formulate:

$$f(i) = \min_{j \in J}\{c(j,i) + f(j)\}; \quad i \in I \tag{1}$$

$$f\left(i^*\right) = \min_{j \in J}\left\{f\left(i^*\right)\right\}; \quad i \in I \tag{2}$$

where $c(j,i)$ is the weight or distance between vertex j and vertex i. To better compare with the Dijkstra's algorithm, in this study, the forward dynamic programming approach will be used to solve the shortest path problem. Equations (1) and (2) implies that the shortest path from the initial vertices to a new vertex i^* is found, thus the set J and set I should be updated as

$$J = J \cup \left\{i^*\right\}$$

$$I = \{N \not\subset J\}$$

Repeat the above steps until the set I is empty, i.e. $I = \varnothing$. Then stop, the one source shortest path problem has been solved, which means that the shortest paths from the started vertex to all other vertex in the network are found.

2.3 Numerical Example

This section we use a numerical example to illustrate the disadvantage of Bellman's dynamic programming for the source shortest path problem. The problem is to find all the shortest paths to vertex 1. The forward Bellman's dynamic programming

approach presented in Sect. 2.2 would be adopted. It is implemented using the adjacency matrix where the connectivity between nodes remains unknown, this enables he applicability to a large network.

According to the presented network, the adjacency matrix is calculated as

$$c = \begin{bmatrix} 0 & 1 & 3 & \text{inf} & \text{inf} & \text{inf} \\ \text{inf} & 0 & 2 & 7 & 4 & \text{inf} \\ \text{inf} & \text{inf} & 0 & 2 & 3 & \text{inf} \\ \text{inf} & \text{inf} & \text{inf} & 0 & 3 & 5 \\ \text{inf} & \text{inf} & \text{inf} & \text{inf} & 0 & 2 \\ \text{inf} & \text{inf} & \text{inf} & \text{inf} & \text{inf} & 0 \end{bmatrix}$$

where "inf" in the matrix implies that the distance/weights between two vertices is infinite, i.e., there is no direct arc connect them.

The recursive calculation of the dynamic programming in solving the problem can be found as following:

- First step: start from the start vertex, let

$$f(1) = 0;\ J = \{1\};\ I = \{2, 3, 4, 5, 6\}$$

- Second step: calculate the shortest path from vertex in J to the vertex in I. According to Eq. (1) we have

$$\begin{aligned} f\left(i^*\right) &= \min_{j \in J}\{c(j, i) + f(j)\}; \quad i \in I \\ &= \min[c(1, 2) + f(1), c(1, 3) + f(1), c(1, 4) \\ &\quad + f(1), c(1, 5) + f(1), c(1, 6) + f(1)] \\ &= \min[1, 3, +\infty; +\infty; +\infty] = 1 \end{aligned}$$

Thus we have $i^* = 2;\ J = \{1, 2\};\ I = \{3, 4, 5, 6\}$ and $j^*(i^*) = 1$;

- Third step:

$$\begin{aligned} f\left(i^*\right) &= \min_{j \in J}\{c(j, i) + f(j)\}; \quad i \in I \\ &= \min\begin{bmatrix} c(1, 3) + f(1), c(1, 4) + f(1), c(1, 5) + f(1), c(1, 6) + f(1), \cdots \\ c(2, 3) + f(2), c(2, 4) + f(2), c(2, 5) + f(2), c(2, 6) + f(2) \end{bmatrix} \\ &= \min[3, +\infty; +\infty; +\infty, 3, 8, 5, +\infty] \end{aligned}$$

Here we find two minimum values. Pick one of them randomly, for example the first minimum value, so we have

$$i^* = 3;\ f(3) = 3;\ J = \{1, 2, 3\};\ I = \{4, 5, 6\} \text{ and } j^*(i) = \{1, 2\}$$

- Fourth step:

$$
\begin{aligned}
f\left(i^*\right) &= \min_{j\in J}\{c(j,i)+f(j)\};\quad i\in I\\
&= \min\begin{bmatrix} c(1,4)+f(1), c(1,5)+f(1), c(1,6)+f(1), \cdots \\ c(2,4)+f(2), c(2,5)+f(2), c(2,6)+f(2), \cdots \\ c(3,4)+f(3), c(3,5)+f(3), c(3,6)+f(3), \cdots \end{bmatrix}\\
&= \min[+\infty; +\infty; +\infty, 8, 5, +\infty, 5, 6, +\infty]
\end{aligned}
$$

Similarity, we also have two minimum value in this route, choose the second minimum value, yields

$$i^* = 4;\ f(4) = 5;\ J = \{1,2,3,4\};\ I = \{5,6\} \text{ and } j^*(i^*) = \{3\}$$

- Fifth step:

$$
\begin{aligned}
f\left(i^*\right) &= \min_{j\in J}\{c(j,i)+f(j)\};\quad i\in I\\
&= \min\begin{bmatrix} c(1,5)+f(1), c(1,6)+f(1), c(2,5)+f(2), c(2,6)+f(2), \\ c(3,5)+f(3), c(3,6)+f(3), \cdots \\ c(4,5)+f(4), c(4,6)+f(4) \end{bmatrix}\\
&= \min[+\infty; +\infty; 5, +\infty, +\infty, 6, +\infty, +\infty, 8, 11]
\end{aligned}
$$

Very obvious, we have $i^* = 5$, $f(5) = 5$; $J = \{1,2,3,4,5\}$, $I = \{6\}$ and $j^*(i^*) = \{3\}$

- Sixth step:

$$
\begin{aligned}
f\left(i^*\right) &= \min_{j\in J}\{c(j,i)+f(j)\};\quad i\in I\\
&= \min[c(1,6)+f(1), c(2,6)+f(2), c(3,6)+f(3), c(4,6)\\
&\quad + f(4), c(5,6)+f(5)]\\
&= \min[+\infty; +\infty; +\infty, 10, 7]\\
&= 7
\end{aligned}
$$

Thus, we have $i^* = 5$, $f(6) = 7$; $J = \{1,2,3,4,5,6\}$, $I = \varnothing$, $j^*(i^*) = \{5\}$

Since $I = \varnothing$, the iteration terminated. Table 1 summarize of main results of above calculation in each iteration. From Table 1 we can easily find the shortest from vertex 1 to any other vertices. For example, the shortest path from vertex 1 to vertex 6 is 7, and by tracking $j^*(i^*)$, the direction of the shortest path is $1 \to 2 \to 5 \to 6$.

While the dynamic programming can solve the one source shortest problem, by observing the calculations in each Step, however, the method is very inefficient and computational expensive. It involves many unnecessary calculations are operated. For example, in the second step calculation, the results of $c(1,3) + f(1), c(1,4) +$

Table 1 Summary of the calculations

Step	i^*	$f(i^*)$	$j^*(i^*)$
1	1	0	∅
2	2	1	{1}
3	3	3	{1, 2}
4	5	5	{2}
5	4	5	{3}
6	6	7	{5}

$f(1), c(1,5)+f(1), c(1,6)+f(1)$ have already obtained, but in the third step, it re-calculates them, this is highly undesirable since it adds the computation effort which is a main indicator of the efficiency of the method. As can be seen, to solve the one resource shortest path problem, the dynamic programming calculated the vertex to vertex shortest path (updata $c(j,i)+f(j)$) for 36 times $(1+5+8+9+8+5)$ which making the time complexity $O(35)$. However most the calculations are repetitive work and can be avoided. The problem occurs because that the dynamic programing has no memory that it cannot record the calculations done previously. If some method that can occupy the deficiency, the efficiency of the dynamic programing would be greatly improve, this is why Dijkstra's algorithms come into play.

2.4 Dijkstra's Algorithm

Dijkstra's algorithm is a graphic search algorithm proposed by Dijkstra [7] to solve the single-source shortest path problem of a graph with non-negative edge path costs. This algorithm widely adopted to find the shortest path in many other graphic algorithms.

Since Dijkstra's algorithm is extensively used by researchers and engineers and is almost found in every textbook, the theoretical proof of effectiveness of this method is omitted. But the procedures to implement the algorithm is presented to illustrate how this method works and to make a comparison with the Bellman's dynamic programming approach.

We denote $c(r, s)$ as the weight of the arc (r, s). Dijkstra's Algorithm marks the vertices as permanent or temporary vertices. We use $P(i)$ and $T(i)$ $i \in N$ to represent the permanent and temporary vertices, respectively, where $P(i)$ denotes the weight of the shortest directed u–r path. A temporary label $T(r)$ gives an upper limit to this weight (can be ∞). Furthermore, we denote:

$$\gamma(r) = \begin{cases} 1, \ if\ the\ label\ is\ permanent \\ 0, \ if\ the\ label\ is\ temporary \end{cases}$$

$$\pi(r) = \begin{cases} if\ the\ predecessor\ of\ vertex\ r\ on\ the\ shortest \\ \quad directed\ u - r\ if\ such\ a\ path\ exists \\ 0, otherwise \end{cases}$$

The step to step procedure for implementing Dijkstra's Algorithm:

Step 1: Set $P(u) \leftarrow 0$ and $\gamma(u) \leftarrow 1, \pi(u) \leftarrow 0$. For all other vertices r, set $T(r) \leftarrow \infty$ and $\gamma(r) \leftarrow 0$. For all vertices r, we set $\pi(r) \leftarrow 0$. Furthermore, set w $\leftarrow$ u
Step 2: For every arc (w, r), where $\gamma(r) = 0$ and $T(r) > T(w) + c(w, r)$, set $T(r) = T(w) + c(w, r)$ and $\pi(r) \leftarrow w$
Step 3: find a vertex r^* for which $\gamma(r^*) = 0$, and $T(r*) = \min_{\gamma(r)=0} \{T(r)\}$, set $\gamma(r^*) = 1$, $P(r^*) = T(r^*)$ and $w \leftarrow r^*$
Step 4: If $w = v$, stop, otherwise go to step #2.

It can be seen from above steps that in each step, not only the permanent labels are updated, but also the T labels are updated at the same time. In this way, the Dijkstra's algorithm avoids the repeat calculations in the Bellman's dynamic programming approach.

2.5 Numerical Example

For comparison, the same numerical example network in Fig. 1 is used to illustrate how the Dijkstra's algorithm can be implemented. Also the problem is to find all the shortest path from all other nodes to nodes 1.

- First Step: let $w = 1$
  Let $P(1) = 0; T(2) = T(3) \cdots T(6) = +\infty; \pi = [\{1\}\ \{0\}\ \{0\}\ \{0\}\ \{0\}\ \{0\}]$;
  $\gamma = [1\ 0\ 0\ 0\ 0\ 0]$;
  After the first round updating T and π in Table 2, we find that $\min_{\gamma(r)=0} \{T(r)\} = T(1)$.
  Therefore, we have $P(2) = 1; \gamma(2) = 1; w = 2$. $\pi = [\{1\}\ \{1\}\ \{1\}\ \{0\}\ \{0\}\ \{0\}]$
- Second Step:

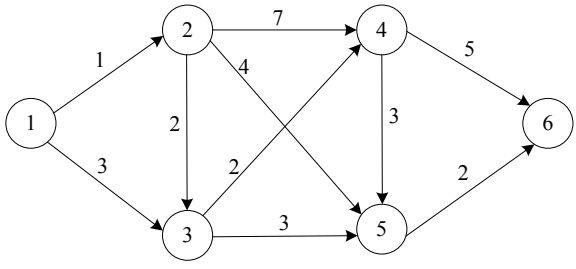


Fig. 1 A numerical example network

Table 2 Summation of the calculations in step 2

Node (i)	T(i) = min[T(i), c(1, i) + $P(1)$]	Update $\pi(r)$
2	T(2) = min[+∞, 1 + 0] = 1	1
3	T(3) = min[+∞, 3 + 0] = 3	1
4	T(4) = min[+∞, +∞ + 0] = +∞	–
5	T(5) = min[+∞, +∞ + 0] = +∞	–
6	T(6) = min[+∞, +∞ + 0] = +∞	–

Table 3 Summation of the calculate in second step

Node (i)	Update T(i), T(i) = min[T(i), c(2, i) + $P(2)$]	Update $\pi(r)$
3	T(3) = min[3, 2 + 1] = 3	1, 2
4	T(4) = min[+∞,7 + 1] = 8	2
5	T(5) = min[+∞,4 + 1] = 5	2
6	T(6) = min[+∞, +∞ + 1] = +∞	–

Now we got a new node with permanent label, i.e., node 2. The temporary vertices (these still marked by T value) are vertices 3, 4, 5, 6, 7, start from node 2, update their T values as shown in Table 3.

Based on the calculation in Table 3, we find $\min_{\gamma(r)=0}\{T(r)\} = T(3) = 3$. Therefore, $P(3) = 3$; $\gamma(3) = 1$, $w = 3$, π = [{1} {1} {1, 2} {2} {2} {0}]

- Third step:
  In step 2, we got a new node (node 3) with permanent label. The temporary vertices (these which are still marked by T value) are vertices 4, 5, 6, 7, then start from node 3, the T values are updated as shown in Table 4.
  Based on the calculations in Table 4, we have $\min_{\gamma(r)=0}\{T(r)\} = T(4)$ *or* $T(5) = 5$; Since the minimum T values are more than 1, we randomly take one of them, assume $\min_{\gamma(r)=0}\{T(r)\} = T(4)$, thus there have
  $P(4) = 5$, $\gamma(4) = 1$, $w = 4$, π = [{1} {1} {1, 2} {3} {2} {0}];
- Fourth step:
  In step 3, node 4 is marked with permanent label. The remaining node with T label is 6, 7. Then start from node 4, the T labels are updated as shown in Table 5.
  Based on Table 5 calculation, we find $\min_{\gamma(r)=0}\{T(r)\} = T(5) = 5$; therefore, we have $P(5) = 5$; $\gamma(5) = 1$; $w = 5$. $\pi = \left[\{1\}\ \{1\}\ \{1, 2\}\ \{3\}\ \{2\}\ \{4\} \right]$.

Table 4 Summation of the calculate in third step

Node (i)	Update T(i), T(i) = min[T(i), c(3, i) + $P(3)$]	Update $\pi(r)$
4	T(4) = min[8, 2 + 3] = 5	3
5	T(5) = min[5, 3 + 3] = 5	–
6	T(6) = min[+∞, +∞ + 3] = +∞	–

Table 5 Summation of the calculate in fourth step

Node (i)	Update T(i), T(i) = min[T(i), c(4, i) + $P(4)$]	Update $\pi(r)$
5	T(5) = min[5, 3 + 5] = 5	–
6	T(6) = min[+∞, 5 + 5] = 10	4

Table 6 Summation of the calculate in fifth step

Node (i)	Update T(i), T(i) = min[T(i), c(5, i) + $P(5)$]	Update $\pi(r)$
6	T(6) = min[10, 2 + 5] = 7	5

- Fifth step:
  The temporary vertices are vertices 5, 6, 7; Then start from newly added node with permanent label 4, update their T values as shown in Table 6:
  Based on Table 6 calculation, we find $\min_{\gamma(r)=0}\{T(r)\} = T(6) = 7$; therefore, $P(6) = 7$; $\gamma(6) = 1$; $w = 6$. $\pi = [\{1\}\ \{1\}\ \{1, 2\}\ \{3\}\ \{2\}\ \{5\}]$; Since $w = v$, thus the calculation is stopped.

According to Table 6 calculations, it can be seen that the shortest path from node 1 to node 6 is 7, and based on the vertices adjacent set π, the shortest path is found to be $1 \rightarrow 2 \rightarrow 5 \rightarrow 6$, which is exactly as it is obtained by the Bellman's dynamic programming method. The shortest path from node 1 to other nodes can also be found in a similar way. It should notes that to find the shortest path, Dijkstra algorithm just update the T value for $5 + 4 + 3 + 2 + 1 = 15$ times, which is less than it is of the dynamic programming method (i.e., 36 times). Thus the Dijkstra's algorithm can greatly improve the efficiency of the dynamic programming.

2.6 Application of the Two Shortest Paths in Large Road Network

To further compare the two algorithms, the real road network in Fig. 2 are used for testing. Both of the networks are well known network in network modeling domain and is often used for testing traffic assignment algorithms. The first network, titled Sioux Fall network, is a medium size network with 24 nodes and 76 link, while the second network, named Borman network, is a large network with 197 nodes and 485 links. The dynamic programming algorithm and Dijkstra's algorithm are programed using Matlab in a computer with 2.1 Hz CPU, 2 GB memory. The results are presented in Table 7. It can be seen that to solve the one source shortest problem on Sioux fall network, the Dijkstra's algorithm just take half of the time of the Bellman's dynamic programming approach. It updates the weights with Eq. (1) for 276 times, much less than it is of dynamic programming which is 2300 times. For the Borman network, the times for updating the shortest path for dynamic programming increases to more

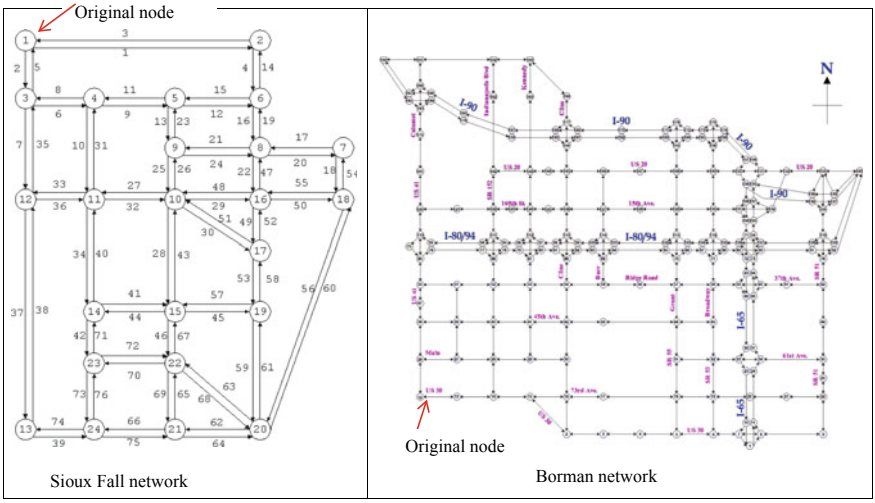


Fig. 2 Two large networks for testing shortest path algorithms

Table 7 Comparison of the performance of the two algorithms in solving the single source shortest path problem in large networks

Network	Number of nodes	Number of links	Origin node	Bellman's Dynamic programming		Dijkstra' algorithms	
				Count	CPU time (s)	Count	CPU time (s)
Sioux Fall	24	76	1	2300	0.003200	276	0.001638
Borman	197	485	12	1,274,196	0.514779	19,306	0.009224

than 1 million, but the Dijkstra's just update them for less than 20 thousand, thus solving the problem much more quickly than dynamic programming.

3 Dynamic Programming for Solving K Shortest Path Problem

3.1 Dynamic Programing Approach for Solving K Shortest Path Problem

As is presented in the introduction part, in the real world, sometime it is difficult to find the real shortest path because of the difficulty to precisely measure the weights of distance of the arcs/links. Sometimes if the bridges or some main road in the shortest path is blocked because of congestion or accidents, then we may want to search for the second shortest path, and if the second shortest path can not be used, we would go the third one. That is K shortest path come into play. Literally, K shortest path studies the paths whose length of total weights is among the minimum, but K shortest path can also be used to refer as the paths whose difference in travel time or distance with the best one's is within the predetermined threshold. In reality, not all drivers take the shortest path because of lack of knowledge or personal preferences, but typically, the routes they choose is among the best ones. The significance of K shortest path is that it can help to find which path the drives would choose thus can give some directions or ideas for the transportation planner to improve the road network.

Similar to shortest path, the K shortest path can also be solved by dynamic programming. Assume a road network G(N, A), where N is number of nodes and A is number of links in the road network. The K shortest path is to find all the paths that satisfy some conditions between the original node u, $u \in A$ and the destination node v, $v \in A$. Let $c(x, y)$ denotes the length of arc (x, y), and $f(x)$ denotes the shortest path from node x to destination node v. $f(x)$ can be obtained by using the once source shortest path algorithms like the dynamic programming and Dijkstra's algorithm introduced previously. According to the methods developed by Byers and Waterman [4], the arc is said to enter if

$$d + c(x, y) + f(y) \leq f(u) + e \tag{3}$$

d = the distance from the original node u to the node x.
e = preset threshold it could be a distance which denotes the path should be within e of the shortest path.
$f(y)$ = shortest path from node y to destination node.

Equation (3) can be taken as the transition equation in finding the K shortest path. The approach uses three attributes:

x a next to last node;
y a last node;
t the length of the path $(1, \ldots, x, y)$ having a subpath $(1, \ldots, x)$ in P, where P is feasible path.

The approach is as following:

Step 1. Set P = (1) and $x = 1$. Then for each arc (1, y) that satisfies $c(1, y) + f(y) \leq f(1) + e$, create an entry $(1, y, t)$ in the stack with $t = \mathrm{c}(1, y)$
Step 2. Stop if the sack is empty
Step 3. Remove (POP) the topmost entry (x, y, t) in the stack. Replace P = $(1, \ldots, x)$ by $(1, \ldots, x, y)$. If $y =$ end node, go to step 4. If not, let $x \leftarrow y$ *and* $d \leftarrow t$. Then for each arc (x, y) satisfying (2), create an entry (x, y, t) in the stack with $t = d + \mathrm{c}(x, y)$. Go to step 2.
Step 4. Output *P* and *t*. Go to step 3.

3.2 Numerical Application

The following we will illustrate in detail how to use the proposed dynamic programming method to solve K shortest path problem. The same small example network in Fig. 1 is used, but to make it clearer, the node mark $1, 2, \cdots$ is replaced by $A, B, \cdots$. The problem is to find the path from node A to node F that is within 1 of the shortest path (i.e. $e = 1$). From previous example, we find that the minimum cost from A to F is 7. Thus, the problem can be changed into searching for the paths with cost that is equal or less than 8. Based on the approach, the minimum cost from each node to the destination node which is node F should be calculated in advance. This could be attained by operating the backward method based on dynamic programming. The minimum cost from each node to the node F is presented in Fig. 4 which is highlighted in blue.

The summary of the calculations by using dynamic programming to find the K shortest path is presented in Table 8. Firstly, we start from node A, since $0+c(A, B)+f(B) = 7 \leq 8$ and $0+c(A, C)+f(C) = 8 \leq 8$ then arc (A, B) and (A, C) enter into the pack. Then we turn to node B since it is newly added node, there are three nodes that the node B direct to, but only node E that satisfy $1 + c(B, E) + f(E) = 7 \leq 8$, thus node E enter into the stack. For node E there just one node connecting with it,

Table 8 Summary of the calculations of using dynamic programming to find the K shortest path

Step	Entries			Path *S*
	Vertex *x*	Vertex *y*	*d*	
1	A A	B C	1 3	(A, B)
2	B A	E C	5 3	(A, B)
3	E A	F C	7 3	(A, B, E)
4	A	C	3	Output (A, B, E, F), $d = 7$
5	C	E	6	(A, C, E)
6	E	F	8	(A, C, E, F), $d = 8$

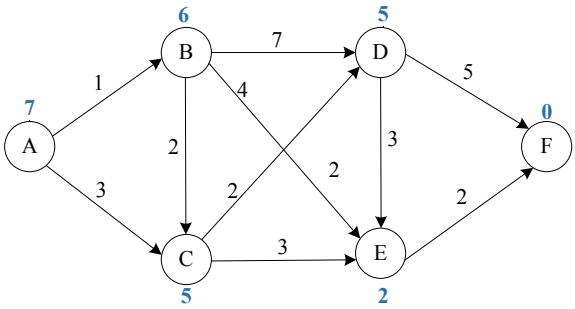


Fig. 3 Numerical example network

i.e. the last node F, since $5 + c(E, F) + f(F) = 7 \leq 8$ thus node F enters. Note that since the destination node is found, then we output the path and cost of the path which is (A, B, E, F) and $d = 7$ respectively. Then in the pack there just arc (A, C) is left, by operating the same procedures, we can also find another feasible path satisfying the preset condition, which is path (A, C, E, F), and the corresponding cost is 8.

The example network in Fig. 3 is just a small example for purpose of illustration. The following we will explore whether the algorithm can be applied in large network. The network topology is presented in Fig. 4. The network includes 64 nodes and 112 links. The original node is node 1, and destination node is node 64. By using the Dijkstra's algorithm, we find two shortest paths between node 1 and node 64, the cost of both of the path is 50. The routes of the two shortest paths are listed as following,

$$\left\{\begin{array}{l} 1 \rightarrow 9 \rightarrow 10 \rightarrow 18 \rightarrow 19 \rightarrow 20 \rightarrow 28 \rightarrow 29 \rightarrow 30 \rightarrow 38 \rightarrow 39 \\ \quad \rightarrow 47 \rightarrow 48 \rightarrow 56 \rightarrow 64 \\ 1 \rightarrow 9 \rightarrow 10 \rightarrow 18 \rightarrow 19 \rightarrow 20 \rightarrow 28 \rightarrow 29 \rightarrow 30 \rightarrow 38 \rightarrow 39 \\ \quad \rightarrow 40 \rightarrow 48 \rightarrow 56 \rightarrow 64 \end{array}\right\}$$

We then want to use the proposed dynamic programming approach to find the K shortest path problem. The threshold is that the path should be within 3 of the minimum cost route, i.e. $e = 3$ in Eq. (2). Since the network in Fig. 3 is a very larger network, we can not calculate manually, thus the approach is programmed with Matlab to find the feasible path. The code is attached in the appendix. The approach took less than 0.05 s to find 4 more paths from node 1 to node 4 which satisfy the threshold. Thus, we find that the dynamic programming approach can efficiently solving the K shortest path problem. The four more paths found by dynamic programming approach is presented in Table 9 and corresponding direction is drawn in Fig. 4.

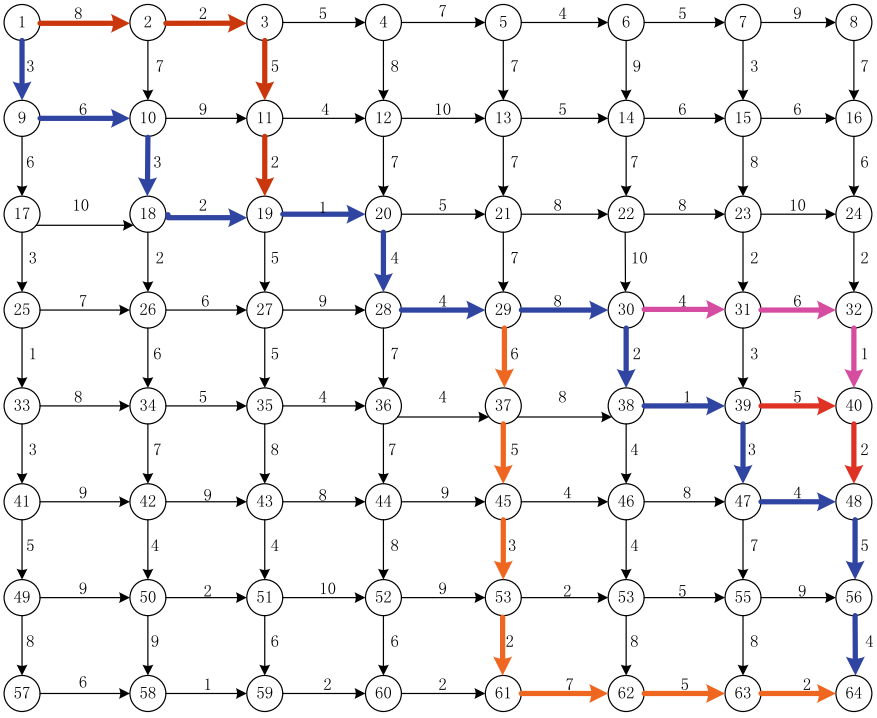


Fig. 4 Numerical network for K shortest path problem

Table 9 Four more paths find by dynamic programming for K shortest problem

Paths	Path directions	Cost
1	1 → 2→3 → 11 → 19 → 20 → 28 → 29 → 30 → 38 → 39 → 47 → 48 → 56 → 64	53
2	1 → 2→3 → 11 → 19 → 20 → 28 → 29 → 30 → 38 → 39 → 40 → 48 → 56 → 64	53
3	1 → 9 → 10 → 18 → 19 → 20 → 28 → 29 → 37 → 45 → 53 → 61 → 62 → 63 → 64	53
4	1 → 9 → 10 → 18 → 19 → 20 → 28 → 29 → 31 → 32 → 40 → 48 → 56 → 64	53

4 Conclusion

This paper studies the shortest path problem. Two algorithms including Bellman's dynamic programming approach and Dijkstra's algorithm are presented. Application of the dynamic programming approach in a small network finds that in each iteration, it repeatedly updates the shortest paths of some nodes which have already done previously. This is highly undesirable since it add the computation effort to solve the shortest problem. The problem occurs because the dynamic programming approach has no "memory" that it cannot save the calculation results done in previous iterations. The Dijkstra's algorithm overcomes this deficiency by marking each node with T label or P label, and each time P label is updated according to the T label which

can save previous best results. The application of the two algorithms in a medium size and a large size road network finds that the Dijkstra's algorithm solves the shortest path problem much quickly than dynamic programming does. This paper also explores applicability of dynamic programming in solving K shortest path problem. An approach based on Byers and Waterman's research is presented. In addition, step to step procedures for applying the approach are illustrated through solving K shortest paths in a small network. The subsequent application of the approach finds that it is also very efficiency in finding the K shortest paths in a large network thus is applicable to solve the real world K-shortest path problem.

Acknowledgements and Conflict of Interest Statement This research is supported by the Natural Science Foundation of Zhejiang province (LQ17E080007), National Natural Science Foundation of China (71501009) and (51408323) and Sponsored by K. C. Wong Magna Fund in Ningbo University. The authors are very grateful to the anonymous reviewers for their valuable suggestions and comments. The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Appendix

Matlab code for dynamic programming approach and Dijkstra's algorithm to solving the shortest path problem on Sioux Fall network and Borman network.

```
clc
clear
data=textread('Cost.txt'); % here the cost in the adjacent matrix of
the network, since the adjacent matrix for Sioux fall and Borman
network is very large. I will not attach them here, but they can be
find the website: http://www.bgu.ac.il/~bargera/tntp/. This network has
many transportation network data for testing.
data(:,3)=data(:,3)./(333.333);
m=max(data(:,1));
A=inf(m,m);
for i=1:length(data)
    A(data(i,1),data(i,2))=data(i,3);
end

s=1;
e=20;
tic
[Count1,distance1,Distance1]=DynamicP(A,s,e); %using dynamic
programming to solve the single source shortest path problem
toc
s=20;
e=1;
tic
[Count1,distance1,Distance]=dijkstra(A,s,e); %using dynamic programming
to solve the single source shortest path problem
Toc


% Dynamic programming approach
function [Count,distance1,Distance]=DynamicP(A,s,e)
% [Count1,distance1,Distance]=DynamicP(A,s,e)
% returns the distance and path between the start node and the end
node.
%count: the number of times for updating shortest path.
% distance1: the shortest length from node s to all other nodes in the
network.
% Distance: the shortest path from node s to node e
% A: adjcent matrix
% s: start node
% e: end node
% initialize
m=size(A);
pb(1:length(A))=0; %indicator, to see wheher all nodes are calculated.
pb(e)=1;
T(1:length(A))=inf;
temp=s;
distance=inf(1,m);
distance(e)=0;
parent=zeros(1,m);
Count=0;
while sum(pb)<length(A)
    index1=find(pb==0);  % find out the node that have not been marked
    index2=find(distance~=inf); % find out the node that have been
marked
    for i=1:length(index1)
        for j=1:length(index2)
           DT(i,j)=distance(index2(j))+A(index1(i),index2(j));
           Count=Count+1;
        end
    end
    [mm,n]=find(DT==min(min(DT)));
    distance(index1(mm(1)))=min(min(DT));
```

```
    parent(index1(mm(1)))=index2(n(1));
    pb(index1(mm(1)))=1;
    DT=[];
end

Distance=distance(s);
% to find the exact shortest path from node S to node E
path=zeros(1,2*m);      % path preallocation
t=s; path(1)=t; count=1;
while t~=e && t>0
    p=parent(t);
    path=[path(1:count) p];
    t=p;
    count=count+1;
end
if count>=2*m
    error(['The path preallocation length is too short.',...
        'Please redefine path preallocation parameter.']);
end
path(1)=s;
path=path(1:count);


%dijkstra's algorithms
function [Count,distance, distance]=dijkstra(A,s,e)
% [Count,distance1,Distance]=dijkstra(A,s,e)
% returns the distance and path between the start node and the end
node.
%count: the number of times for updating shortest path.
% distance1: the shortest length from node s to all other nodes in the
network.
% Distance: the shortest path from node s to node e
% A: adjcent matrix
% s: start node
% e: end node
% initialize
m=size(A);
% s=1;
% e=3;
pb(1:length(A))=0;pb(s)=1;
T(1:length(A))=inf;d(1)=0;temp=s;
distance=inf(1,m);
distance(s)=0;
parent=zeros(1,m);
Count=0;
while sum(pb)<length(A)
    tb=find(pb==0);
    for i=1:length(tb)
        if T(tb(i))>distance(temp)+A(temp,tb(i))
            T(tb(i))=distance(temp)+A(temp,tb(i)); % updata the T mark
            parent(tb(i))=temp;
            Count=Count+1;
        end
    end
    aa=min(T(tb));
    tmpb=find(T(tb)==aa);
    temp1=tb(tmpb(1));
    distance(temp1)=aa;
%     parent(temp1)=temp; % updata the P mark
    temp=temp1;
    pb(temp)=1;
end

Distance=distance(e);
```

Matlab code for solving the K shortest path problem

```
clc
clear

data=load('A','A');
A=data.A; % A is adjacent matrix, because it is very large for network
in Figure 4. I will not attach here

s=1;
e=64;
[distance,Distance,path1]=DynamicP(A,s,e); %to solve all the shortest
path to the destination node numbered 64.

Td=Distance+3;
s=1;
e=64;
path=zeros(200,24); % to store the path that satisfy the threshold
result=[];
index_result=1;
n=0;
for j=1:length(A)
    if A(s,j)+distance(j)<=Td & s~=j
        n=n+1;
        path(n,1:2)=[s,j];
        path(n,24)=A(s,j);
    end
end

index2=0;
index2=index2+n;;  % to record the number of path in the stack
while path(1,1)~=0
    index=find(path(1,:)~=0);
    start=path(1,index(length(index)-1));
    Md=path(1,1:length(index)-1);
    Kshort=path(1,24);
    ss=0;          % to indicate how many sub-nodes that in the shortest
path.
    for j=1:length(A)
        if A(start,j)+Kshort+distance(j)<=Td & start~=j
            ss=ss+1;
            if ss==1
                ss1=1; % the code here is to show whether feasible node
find is the first one, if it is, then the added it to the first row,
else write it in a new row
                ss2=0;
            else
                ss1=2;
                ss2=1;
            end
            path(ss1:index2+ss2,:)=path(1:index2,:);
            path(1,1:length(index))=[Md j];
            path(1,24)=A(start,j)+Kshort;
            index2=index2+1;
            Ed_indicator=j;
            if Ed_indicator==e
                result(index_result,:)=path(1,:);
                path(1,:)=[];
                index2=index2-1;
                index_result=index_result+1;
            end
        end
    end
end
```

References

1. Aljazzar H, Leue SK (2011) A heuristic search algorithm for finding the k shortest paths. Artif Intell 175(18):2129–2154
2. Bauer R, Delling D (2009) SHARC: fast and robust unidirectional routing. ACM J Exp Algorithmics 14(2–4):1–29
3. Bellman R (1958) On a routing problem. Q Appl Math 16:87–90
4. Byers TH, Waterman MS (1984) Technical note—determining all optimal and near-optimal solutions when solving shortest path problems by dynamic programming. Oper Res 32(6):1381–1384
5. Delling D, Goldberg AV, Pajor T, Werneck RF (2013) Customizable route planning in road networks. Announced at SEA 2011 and SEA 2013, full version available online at http://research.microsoft.com/pubs/198358/crp_web_130724.pdf
6. Denardo EV, Fox BL (1979) Shortest-route methods: 1. Reaching, pruning, and buckets. Oper Res 27:161–186
7. Dijkstra EW (1959) A note on two problems in connexion with graphs. Numer Math 1(1):269–271
8. Eppstein D (1998) Finding the k shortest paths. In: 35th IEEE symposium. SIAM J Comput 28(2):652–673
9. Feillet D, Dejax P, Gendreau M, Gueguen C (2004) An exact algorithm for the elementary shortest path problem with resource constraints: application to some vehicle routing problems. Networks 44:216–229
10. Floyd Robert W (1962) Algorithm 97: shortest path. Commun ACM 5(6):345
11. Goldberg AV (2001) Shortest path algorithms: engineering aspects. In: ISAAC 2001, pp 502–513
12. Güner AR, Murat A, Chinnam RB (2009) Dynamic routing using real-time ITS information. Intelligent Vehicle Controls and Intelligent Transportation Systems—Proceedings of the 3rd International Workshop
13. Güner AR, Murat A, Chinnam RB (2012) Dynamic routing under recurrent and non-recurrent congestion using real-time ITS information. Comput Oper Res 39(2):358–373
14. Hart PE, Nilsson NJ, Raphael B (1968) A formal basis for the heuristic determination of minimum cost paths. IEEE Trans Syst Sci Cybern SSC4 4(2):100–107
15. Holzer M, Schulz F, Wagner D (2008) Engineering multi-level overlay graphs for shortest-path queries. ACM JEA 13(2.5):1–26
16. Maue J, Sanders P, Matijevic D (2006) Goal-directed shortest-path queries using precomputed cluster distances. ACM J Exp Algorithmics 4007:316–327
17. Maue J, Sanders P, Matijevic D (2010) Goal-directed shortest-path queries using precomputed cluster distances. J Exp Algorithmics 14:2.3.2–2.3.27
18. Pettie S (2004) A new approach to all-pairs shortest paths on real-weighted graphs. Theoret Comput Sci 312:47–74
19. Schulz F, Wagner D, Weihe K (2000) Dijkstra's algorithm on-line: an empirical case study from public railroad transport. ACM JEA 5(12):1–23
20. Seidel R (1995) On the all-pairs-shortest-path problem in unweighted undirected graphs. J Comput Syst Sci 51(3):400–403
21. Thorup M (1999) Undirected single-source shortest paths with positive integer weights in linear time. J ACM 46(3):362–394
22. Warshall S (1962) A theorem on Boolean matrices. J ACM 9(1):11–12
23. Wilson DB, Zwick U (2013) A forward-backward single-source shortest paths algorithm. In: IEEE 54th annual symposium on foundations of computer science, pp 707–716
24. Yen JY (1971) Finding the K-shortest loopless paths in a network. Manage Sci 17(1):712–716
25. Zhan F (1998) Three fastest shortest path algorithms on real road networks: data structures and procedures. J Geogr Inf Decis Anal 1(1):69–82

Understanding of Day-to-Day Route Choice Behavior: Experiments and Simulations

Lingmin Yang, Rihui She, Jingyi An, Hong Wang and Shunying Zhu

Abstract Dissension arises on whether the day-to-day route choice behavior will cause an equilibrium distribution of traffic flow on the road network, and the travelers' decision-making mechanism of route choice behavior is still in the exploratory stage. This paper focuses on the 'equilibrium dissension' and the 'decision making mechanism' under the condition of historical experience and traffic information by conducting human-computer interaction experiments and multi-agent simulations. The results of experiments support the conclusion that 'no convergence to equilibrium had been found'. Moreover, the simulations with the existing mechanism that perceptions of travel time being the criterion of cognition and the logit discrete choice model being the criterion of route selection support the conclusion as well, it is also found through the simulations that dissension may due to exact treatment of treatment on the probability of discrete choice model. At the same time, comparisons between experiments and simulations found that the existing mechanism was not sufficient to reflect the fact that travelers tend to choose the shorter route more, and the shorter the more when difference of routes' length exist, The modified mechanism proposed by this paper reflect the fact better. This study is beneficial for understanding the traveler's route choice behavior and the causes of traffic congestion.

Keywords Traffic behavior · Day-to-day route choice · Equilibrium · Decision-making mechanism

L. Yang (✉)
College of Management, Hubei University of Education, Wuhan 430205, China
e-mail: yanglingminlily@126.com

R. She
Fuzhou Urban Planning and Design Research Institute, Fuzhou 350003, China

J. An · H. Wang · S. Zhu
School of Transportation, Wuhan University of Technology, Wuhan 430063, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_13

1 Introduction

Traffic congestion is the result of aggregation and distribution of travelers' route choice behaviors on the road network, thus an in-depth understanding of traffic route choice behavior is the key for implementing traffic management measures, and further more easing traffic congestions. The day-to-day route choice behaviors has attracted much attention. On one hand, commuting traffic is the main constituent part of peak hour congestions on work days; on the other hand, the traditional theory of traffic equilibrium has been proved inconformity with reality [1], then how the travelers choose among routes according to historical experience, traffic information and personal judgement, in other words, the traveler's decision mechanism of route choice behavior (or routing mechanism), and whether their choices lead to convergence to equilibrium distribution of traffic flow on the road network has become a problem to discover, the day-to-day route choice behavior provides an excellent perspective to study these problems with its day-to-day evolution characteristics.

However, earlier studies mainly focused on day-to-day dynamics of traffic flow, by assuming traveler' behavioral characteristics to be bounded rationality or to have incomplete information, researches centered on the evolutionary process and mechanism from non-equilibrium state to equilibrium state on the traffic network [2–5]. The essence of related researches is to explore dynamic evolution rules, which guarantee the system to convergence to the Wardrop equilibrium. And under the premise that whether the traffic system will and how to converge to equilibrium, it is difficult to ensure the real-life traffic system work in accordance with these rules.

In fact, whether the day-to-day route choice Behavior will lead to convergence to equilibrium distribution of traffic flow on the road network is in dispute, there are different conclusions in literature [6–8]. Day-to-day route choice evolution has been proven to be "conditional equilibrium" by Guan et al. [6]. With the method of mathematical reasoning and based on the assumption that travelers were individual homogeneity, bounded rationality, and got incomplete information. Liu et al. [7]. proved that the day-to-day route choice could converge to equilibrium by multi-agent simulation, based on the same assumption about travelers' characteristics. Selten et al. [8] conducted laboratory day-to-day route choice experiments with and without traffic information, and draw a conclusion that there is no convergence to equilibrium under both conditions.

Researches on traveler's decision mechanism of route choice behavior (routing mechanism) involves abundant theories and models, such as Prospect Theory [9, 10], Regret Theory [11], Threshold Model [12], Satisfactory Model [13], Learning Model [14], and analysis based on modeling travelers' behavior under historical experiences and traffic information [15, 16]. Although these researches reflect the route choice mechanism based on different conditions and factors from different aspects to a certain extent, due to the complexity of traveler's decision making process and the variety of factors influencing route choice behavior, there is no typical model proven to express traveler's route choice mechanism essentially and veritably, so there is still a long way yet to go for further exploration.

This paper continues with the work of literature [7–9], focuses on the "equilibrium dissension" and the "decision making mechanism" under the condition of historical experience and traffic information by conducting human-computer interaction experiments and multi-agent simulations. Considering the real-life traveler route choice behavior is difficult to organize, this paper alternatively chooses human-computer interaction experiments system based on computer programming. In viewing of the applicability and superiority of the multi-agent modeling and simulation to complex adaptive system characterized with emerging unpredictable overall behaviors from simple behavioral rules of individuals [17], related discussions are made based on simulating traveler's decision-making mechanism with multi agent rules.

2 Human-Computer Interaction Experiments

2.1 The Interaction Systems

The Human-computer interaction experiment system was developed based on Unity 3D and Visual Studio programming, the physical structure of the system is shown in Fig. 1. Participants (Virtual travelers) choose between routes in simulation of actual path scenes at clients, and the server performs backstage computing and data storage.

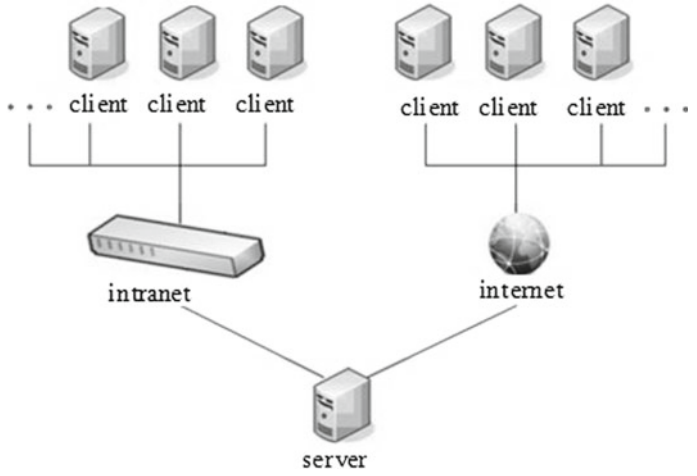


Fig. 1 The physical structure of human computer interaction system

2.2 Experiments Setup

There are 2 routes between O (Origin) and D (Destination), route 1 and route 2, the length of two routes are consider 3 situations of route length between O-D: Situation 1 ($L_1 = L_2$); Situation 2 ($1.5L_1 = L_2$); Situation 3 ($2L_1 = L_2$), details showing in Fig. 2.

The travel time on one route is calculated by the following formula:

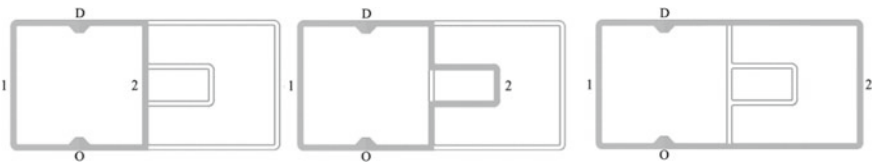
$$t = t_0[1 + \alpha(v/c)^{\varphi}] \tag{1}$$

From above, t is the travel time on one route (route 1 or route 2), t_0 is the travel time of the route under free flow, v is the traffic volume of the route, c is the traffic capacity of the route, α, φ are parameters, $\alpha = 0.15$, $\varphi = 4$.

All the participants are senior students or graduate students majored in traffic engineering, and have conducted enough rounds of trials before the formal experiments, to make sure that the participants are clear about the meaning of Formula (1) and the impact of different number of persons choose path 1 or path 2.

A participant can visually feel the complete process of his/her agent in the interaction experiments going from O to D at each client computer screen. Then the client computer screen will display to the participant the travel times on each route for the last round. Aforementioned measures ensure participants have enough sense of experiencing and the traffic information (here means the travel time) is open and known to all participants.

The similarities and differences between the experiments conducted by this paper and R Selten et,al. are as follow: Selten et al. conducted experiments under two traffic information treatments (In treatment 1, each participant was provided travel time of the last chosen route but not provided travel time of the non-chosen-route in the last round; and in treatment 2, each participant was provided travel times on both routes.) with the length of route 1 and route 2 are equal, and draw a conclusion that there is no convergence to equilibrium under both treatments. Their research did not pay any attention to the impact of length difference on route choice behavior. Considering the circumstances of length difference between the two routes and whether there is consecutive influence on participants' route choice behavior? In this paper, each participant was provided travel times on both routes, and 3 route length difference situations are considered in the experiments.



(a) Situation 1 ($L_1 = L_2$) (b) Situation 2 ($1.5L_1 = L_2$) (c) Situation 3 ($2L_1 = L_2$)

Fig. 2 3 situations of road length between O-D

2.3 Results of Experiments and Analysis

Set the traffic capacity of both route 1 and route 2 to be equal and the exact value being $c = 5$, suppose the number of participants to be n, and x is the number of participants choosing route 1. Thus, the demand of theoretical equilibrium exist under 3 length situations are:

Situation 1:

$$x = 0.5n \tag{2}$$

Situation 2:

$$1 + 0.15\left(\frac{x}{5}\right)^4 = 1.5\left[1 + 0.15\left(\frac{n-x}{5}\right)^4\right] \tag{3}$$

Situation 3:

$$1 + 0.15\left(\frac{x}{5}\right)^4 = 2\left[1 + 0.15\left(\frac{n-x}{5}\right)^4\right] \tag{4}$$

The results of experiments under 3 situations are shown as follows: Fig. 3 for situation 1, Fig. 4 for situation 2, Fig. 5 for situation 3, the corresponding statistics are shown in Table 1. Each situation has been simulated for 30 rounds.

We may see from Figs. 3, 4 and 5 and Table 1 that there is no convergence to theoretical equilibrium under all the 3 situations. Thus, the results of experiments conducted here tend to support the conclusion drawn by Selten et al. [8] that "no tendency of convergence to equilibrium".

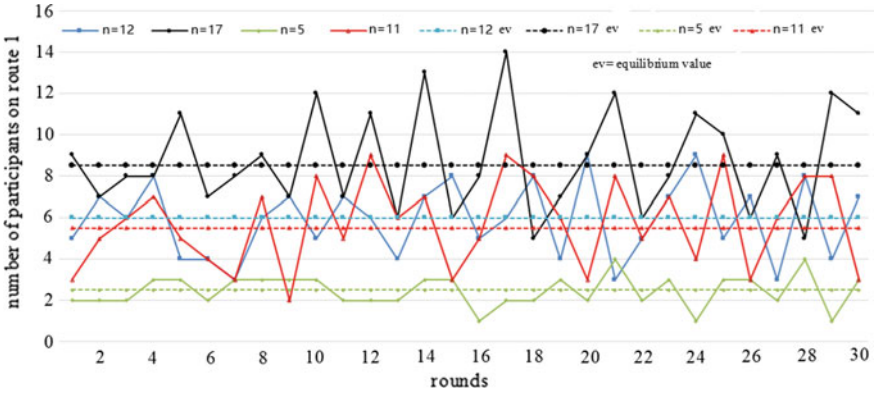


Fig. 3 Theoretical equilibrium value and experimental results under situation 1

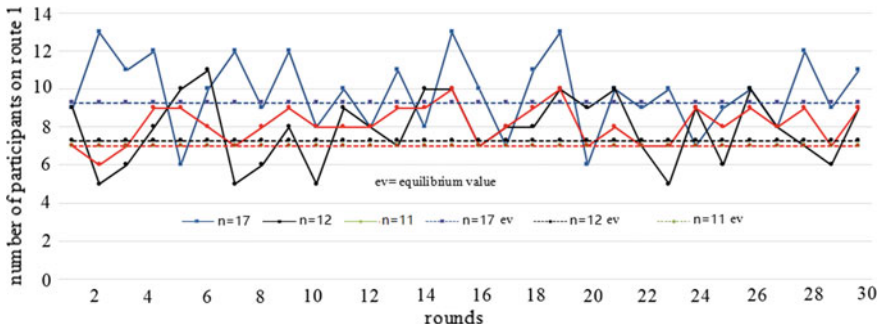


Fig. 4 Theoretical equilibrium value and experimental results under situation 2

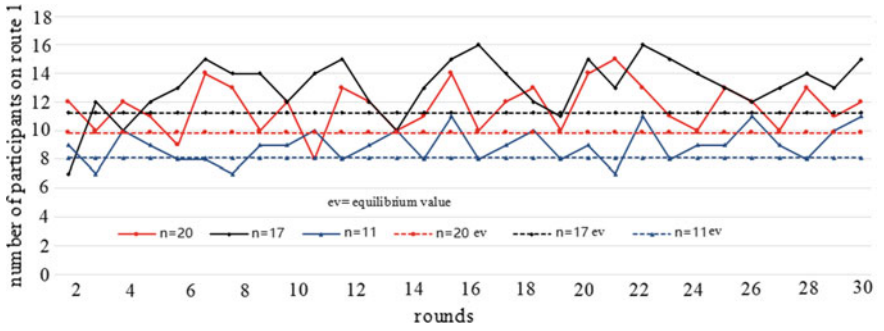


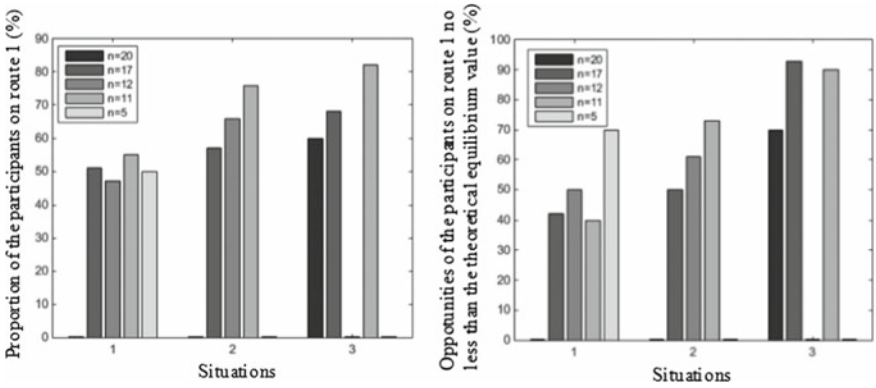
Fig. 5 Theoretical equilibrium value and experimental results under situation 3

Meanwhile, it is worth noting that, with the length of route 1 being invariable and the length of route 2 keeping increasing from situation 1 to situation 3, the proportion of participants on route 1 exceeds the theoretical equilibrium value increasing too. The proportion of participants on route 1 and the opportunities of the participants choosing path 1 no less than the theoretical equilibrium value under 3 situations are shown in Fig. 6.

This shows that travelers tend to choose the shorter length route more when length difference exists, and this phenomenon become more significant when the difference become more obvious, even it resulted in the actual travel time being even longer. This may partly explain why traffic congestions happen often and repeatedly on some paths on the city road traffic network. In addition, at the same time, the parallel paths are comparatively not so crowded. Which is so common in real-life but seems weird according to traffic equilibrium theory, because according to the theory, it is more sensible and costs less travel time for travelers to choose the parallel paths.

Table 1 statistics of experimental results

		Situation 1				Situation 2			Situation 3		
n		17	12	5	11	17	12	11	20	17	11
Equilibrium value on route 1		8.5	6	2.5	5.5	9.3	7.3	7.1	11.2	9.9	8.1
Equilibrium proportion on route 1 (%)		50	50	50	50	55	61	64	56	58	74
Experimental value on route 1	Mean	8.6	5.6	2.5	5	9.7	7.9	8.4	11.9	11.6	9
	Proportion (%)	51	47	50	55	57	66	76	60	68	82
	Variance	1.9	1.4	0.5	2	2.1	1.9	1.1	2.4	1.6	1.2



(a) Proportion of the participants on route 1

(b) Opportunities of the participants on route 1 no less than the theoretical equilibrium value

Fig. 6 Proportion of the participants on route 1 and opportunities of the number of participants on route 1 no less than the theoretical equilibrium value under human-computer experiments

3 Multi-agent Simulation and Analysis

3.1 Simulation Model

Netlogo software is applicable to complex adaptive system modeling and analysis, with its programmable multi-agent modeling function, and has been applied in economic and social researches [18–20]. This paper utilizes this software to fulfil multi-agent modeling and analyzing, the model consists of 2 route (route 1, route 2) agents and multiple experimental participant agents (the number of participant agents are definable by the simulation controller from the input interface), and the participant agents choose between the 2 routes repeatedly.

(1) Rules of route agents

The computing of travel time for route agents is the same with Formula (1), travel time of each route on day i ($i = 1, 2, \ldots, d$) is:

$$t_1^{(i)} = t_{1(0)}\left[1 + 0.15(x^{(i)}/5)^4\right] \tag{5}$$

$$t_2^{(i)} = t_{2(0)}\left[1 + 0.15((n - x^{(i)})/5)^4\right] \tag{6}$$

From above, $t_1^{(i)}$, $t_2^{(i)}$ is separately the travel time on route 1, route 2 on day i, $t_{1(0)}$, $t_{2(0)}$ is separately the travel time on route 1, route 2 under free flow, n is the number of participant agents defined by the simulation controller, $x^{(i)}$ is the number of participant agents on route 1 on day i.

(2) Rules of participant agents

Taking the route choice mechanism proposed by existing researches as rules of participant agents:

Historical experience and traffic information will affect traveler's route choice decisions [15, 16]. Kim et al. assume that the traveler choose between/among paths according to their current perceptions/understanding of travel time on each route, and traveler's perceptions/understanding of travel time on one route is determined by the combination of historical experience and traffic information, which means for participant agents n under traffic information conditions, his perception of travel time for route k ($k = 1, 2$) on day $i + 1$ ($i = 1, 2, \ldots, d$) is:

$$T_{\mathrm{k}}^{(\mathrm{n,i}+1)} = \mu_{\mathrm{k}}^{(\mathrm{n,i}+1)} + \delta_{\mathrm{k}} \tag{7}$$

From above, $\mathrm{T}_{\mathrm{k}}^{(\mathrm{n,i}+1)}$ is the perception of travel time for route k ($k = 1, 2$) on day $i + 1 (i = 1, 2, \ldots, d)$, $\mu_{\mathrm{k}}^{(\mathrm{n,i}+1)}$ is the expectation value of $\mathrm{T}_{\mathrm{k}}^{(\mathrm{n,i}+1)}$, δ_{k} is the random nominal.

According to the assumptions that δ_{k} ($k = 1, 2$) obey the same Gumbel distribution in the logit discrete route choice model, thus, on day $i + 1$, the probability of participant agent n choosing route 1 is:

$$\mathrm{p}_1^{(\mathrm{n,i+1})} = \frac{\exp\left(-\lambda \mu_1^{(\mathrm{n,i+1})}\right)}{\exp\left(-\lambda \mu_1^{(\mathrm{n,i+1})}\right) + \exp\left(-\lambda \mu_2^{(\mathrm{n,i+1})}\right)} \tag{8}$$

From above, λ is the parameter of logit discrete choice model, characterizing the stochastic extent of route choice behavior.

The expected travel time on day $i + 1$ is expressed to be the weighted sum of the actual travel time and the expected travel time on day i [21–23].

$$\mu_{\mathrm{k}}^{(\mathrm{n,i}+1)} = \beta \mathrm{t}_{\mathrm{k}}^{(\mathrm{n,i})} + (1 - \beta)\mu_{\mathrm{k}}^{(\mathrm{n,i})} \tag{9}$$

From above, $\beta \in [0, 1]$, characterizing the trade-off of participant agent n between actual travel time and the expected travel time on day i.

Because of the IID (Independently Identically Distribution, in which all errors have the same variance, resulting in route choice results relating only with the difference value between impedance of two routes) defects, the relative logit route choice model was proposed [24], according to the model, probability of traveler n to choose route 1 on day $i + 1$ is:

$$\mathrm{p}_1^{(\mathrm{n,i+1})} = \frac{\exp(-\varepsilon)}{\exp(-\varepsilon) + \exp\left(-\varepsilon \frac{\mu_2^{i+1}}{\mu_1^{i+1}}\right)} \tag{10}$$

From above, meaning of ε is similar to λ in Formula (5), here use a different symbol for the purpose of discriminating two formula.

3.2 Simulation Results and Analysis

Considering the estimation error of travel time on the first day, experimental simulation was carried out with different initial values (initial number of participants on route 1 keep changing from 1 to n), with the participants being $n = 11$ and $n = 17$.

(1) Influence of values of the parameters

Through simulation comparison, different values of λ have impact on the magnitude of numbers of participants on route 1. With the case $n = 17$, the initial number of participants on route 1 is 11, $\beta = 0.2$, comparison of the number of participants on route 1 is shown in Fig. 7.

When the value of β is relatively small, there is no significant influence observed, but when the value of β is relatively large, the system fall into state of fluctuation. Moreover, it observed that the larger the value of λ being, the system starts to fluctuation since β is smaller. When $n = 17$, the initial number of participants on route is 11, $\lambda = 0.5$, $\beta = 0.2$, comparison of the number of participants on route 1 and 2 in the cases $\beta = 0.1$ and $\beta = 0.7$ is shown in Fig. 8a and b. When $n = 17$, the initial number of participants on route 1 is 11, comparison of the number of participants on route 1 and 2 in the cases $\lambda = 0.5$, $\beta = 0.93$ and $\lambda = 3$, $\beta = 0.75$ is shown in Fig. 8c and d.

In all the subsequent discussions, set $\beta = 0.2$, $\lambda = 0.5$.

(2) Discussion on equilibrium

It is observed in the simulation that there is no convergence to theoretical equilibrium or any other fixed value under various number of participant agents with all possible initial values. There are values of number of participant agents at certain

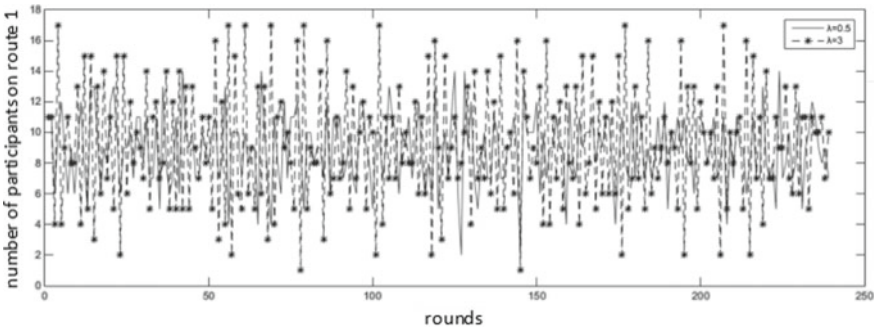


Fig. 7 Influence analysis of parameter λ

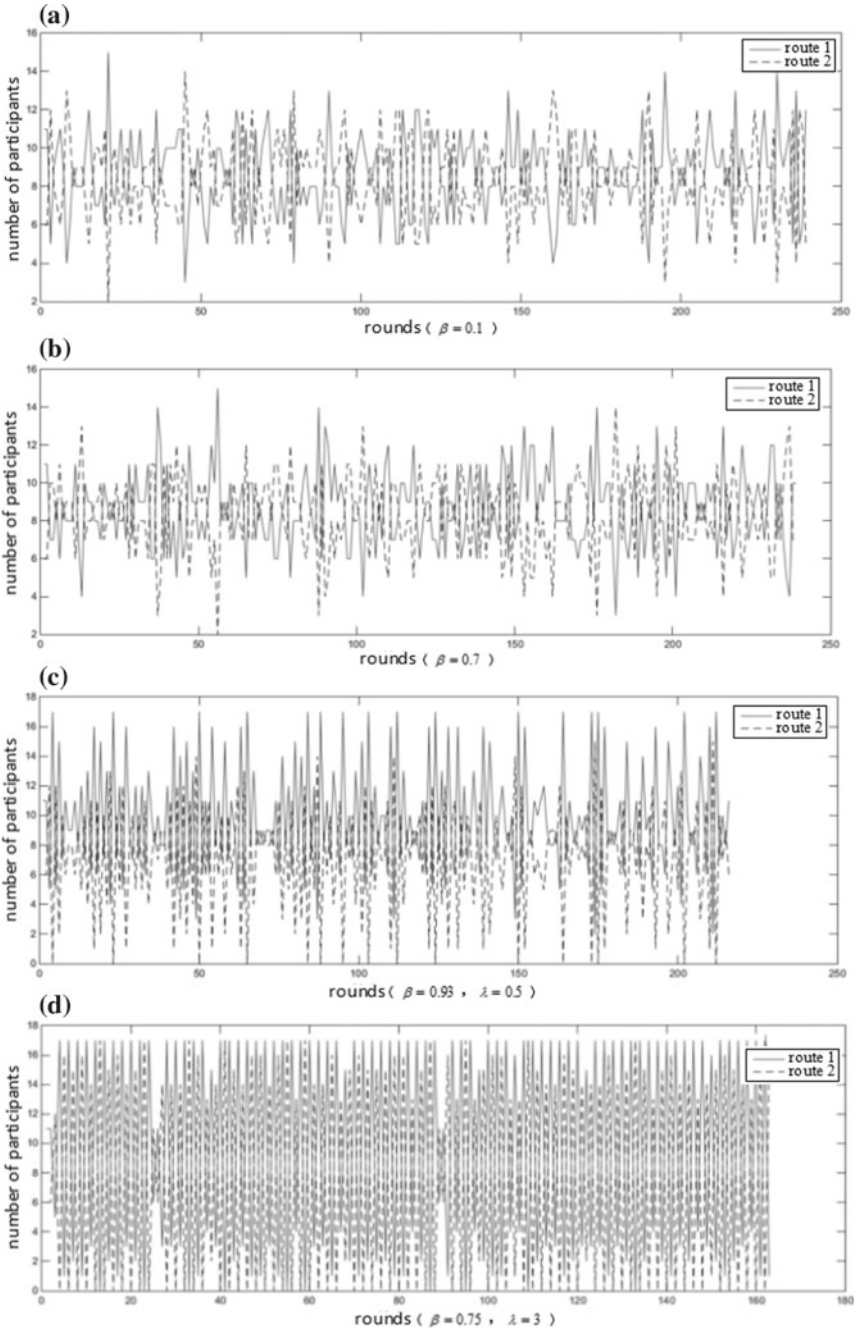


Fig. 8 Influence analysis of parameter β

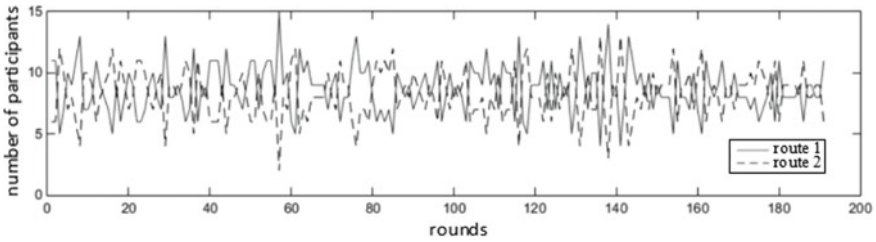


Fig. 9 The changing numbers of participants on 2 routes under multi-agent simulation

moments very close or equal to the theoretical equilibrium value, but soon get into divergence or oscillation. Take situation 2 as an example, when $n = 17$,the initial number of participants on route 1 is 11, the number of participants on route 1 and 2 as time pass by is shown in Fig. 9.

And the results of simulations support the conclusion of "no convergence to equilibrium" which was drawn through results of human-computer interaction experiments.

Considering the conclusion drawn by Guan et,al. that the day-to-day route choice problem is "conditional equilibrium" [6]. The conditions for the existence of evolutionarily stable states are all members of a traveler's group share the same strategy (all members choose route 1 or route 2, these two states are evolutionarily stable states but not equilibrium state) or all members of a traveler's group share the same strategy which ensures a certain proportion of members on each route keep or changing their last choice on the whole (the evolutionarily stable states in this case is equivalent to the equilibrium state under the traffic assignment theory or the stochastic utility theory). And we set the participant agents act according Guan's proposal in literature [6], the results with various values of n with all possible initial values show convergence to theoretical equilibrium, which is similar to Liu's conclusion [7].

The reason for the above circumstances is similar to the probability event of "coin tossing": 10 coins thrown into the air, theoretically, the probability of head upward and tail upward is separately 50%, but the circumstance of 5 coins head upward and other 5 coins tail upward will not always happen for every round, even continuously do the coin tossing for infinitely times, 5 head and 5 tail circumstance may happen rarely. Therefore, the stability condition of the evolution proposed in literature [6] is difficult to achieve in reality. So, the conclusion of no converge to equilibrium observed in reality is not contradictive with Guan et al.'s "conditional equilibrium" conclusion.

Of course, the reason for "no convergence to equilibrium" of the traffic flow in real-life road network may be more than the aforementioned reasons, such as: construction or accident caused temporal capacity change of road paths, the travelers' cognitive actuality of knowing about historical status but not knowing about the future status compared to year by year traffic demand increase and many other factors, which makes it much harder for travelers to deal with in the sense of road traffic status

than descriptions modeled in this paper. Therefore, the traffic flow in real-life road network may even be more difficult to achieve a stable state of equilibrium.

(3) Discussion on existing decision-making mechanisms

Considering the estimation error of travel time on the first day, the average number of participants on route 1 (Ave) and the opportunities of the participants on route 1 no less than the theoretical equilibrium value (Opp) with different initial values (initial number of participants on route 1 keep changing from 1 to n) with the probability calculated through logit discrete route choice model was shown in Table 2 and Table 3. We may find that with different initial values the changes of the average number of participants on route 1 (Ave) and the opportunities of the participants on route 1 no less than the theoretical equilibrium value (Opp) is insignificant. Similar conclusion could be made with the probability calculated through relative logit discrete route choice model, details neglected here. We may think the initial value has little influence on the average number of participants on route 1 (Ave) and the opportunities of the participants on route 1 no less than the theoretical equilibrium value (Opp).

The average number of participants on route 1 and the opportunities of the participants on route 1(Ave) no less than the theoretical equilibrium value (Opp) with the probability calculated through logit model and relative logit model was shown in Tables 4 and 5.

The comparison of average number of participants on route 1 when $n = 17$, $n = 11$, with the probability calculated through logit model, relative-logit model and experimental results was shown in Fig. 10. The comparison of opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value when $n = 17$, $n = 11$, with the probability calculated through logit model, relative-logit model and experimental results was shown in Fig. 11.

We may see that the average number of participants on route 1 and the opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value when $n = 17$, $n = 11$, with the probabilities calculated through logit model and relative-logit model, fit with the experimental results well when the length of route 1 and route 2 is equal, but not fit very well when there is difference between the length of route 1 and route 2.

Therefore, we believe that using the perceived/understanding travel time as criterion of traveler's cognition under the condition of historical experience and open traffic information, with the probability calculated through logit model or relative-logit model, this decision-making mechanism of route choice is not sufficient to reflect the fact that travelers tend to choose the shorter length route more when length difference exists, and this phenomenon become more significant when the difference become more obvious.

Table 2 Average value of the number of participants choosing path 1 (Ave) and opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value (Opp) when n = 11

The initial number of participants on route 1		11	10	9	8	7	6	5	4	3	2	1	0
Situation 1	Ave	5.5	5.5	5.5	5.5	5.5	5.6	5.4	5.5	5.5	5.4	5.5	5.4
	Opp (%)	52	49	50	50	49	53	48	50	49	48	49	50
Situation 2	Ave	6.0	6.1	6.1	6.1	5.9	6.0	6.1	6.0	6.1	6.0	6.0	6.0
	Opp (%)	43	42	44	42	40	41	43	39	42	44	39	40
Situation 3	Ave	6.5	6.5	6.5	6.5	6.5	6.5	6.4	6.4	6.5	6.5	6.5	6.5
	Opp (%)	34	33	35	35	34	32	35	32	35	32	35	35

Table 3 Average value of the number of participants choosing path 1 (Ave) and opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value (Opp) when n = 17

The initial number of participants on route 1		17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Situation 1	Ave	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5
	Opp (%)	50	50	51	50	50	48	50	49	50	50	49	51	50	49	49	49	50	50
Situation 2	Ave	9.1	9.1	9.0	9.1	9.1	9.0	9.1	9.1	9.0	9.0	9.1	9.1	9.0	9.0	9.1	9.0	9.0	9.1
	Opp (%)	58	59	58	59	57	59	58	57	59	59	58	59	58	58	60	57	58	59
Situation 3	Ave	9.5	9.6	9.5	9.5	9.5	9.5	9.6	9.4	9.6	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.4	9.5
	Opp (%)	50	54	50	51	51	52	51	51	54	53	51	53	52	54	53	53	52	52

Table 4 Average value of the number of participants choosing path 1 (Ave) and opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value (Opp) with logit model

Number of participants on route 1	Situation 1		Situation 2		Situation 3	
	n = 17	n = 11	n = 17	n = 11	n = 17	n = 11
Ave	8.5	5.5	9.1	6.1	9.5	6.5
Opp (%)	50	50	58	42	53	34

Table 5 Average value of the number of participants choosing path 1 (Ave) and opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value (Opp) with relative-logit model

Number of participants on route 1	Situation 1		Situation 2		Situation 3	
	n = 17	n = 11	n = 17	n = 11	n = 17	n = 11
Ave	8.5	5.5	9.0	6.1	9.5	7.4
Opp (%)	51	74	32	42	36	53

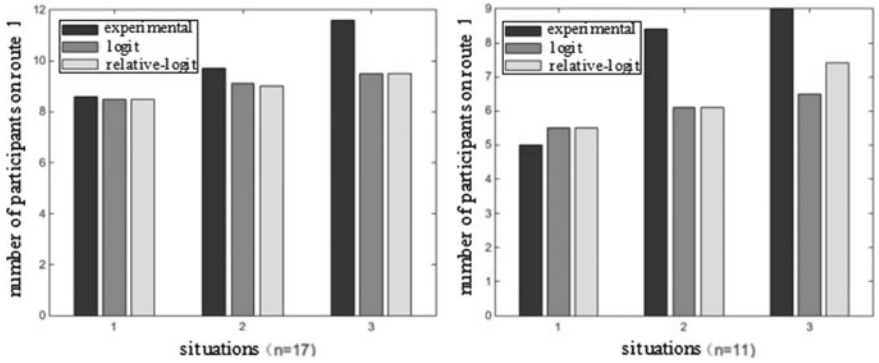


Fig. 10 Comparison of average value of the number of participants choosing path 1(Ave)

4 Discussion on the Traveler's Decision Mechanism

4.1 Criterion of Traveler's Cognition and Influential Factors

Tawfik et al. [25]. found that its not the direct factors like historical experience of a traveler to determine which route to go, but the traveler's perceptions about decision related factors. That is to say, a traveler determines according the "perception" of travel time (perceptive travel time) on choosing between routes. This conclusion confirms the assumption made by Kim et al. [7, 26] to some degree.

The reason for the fact that travelers tend to choose the shorter length route when length difference exists, and this phenomenon become more significant when the

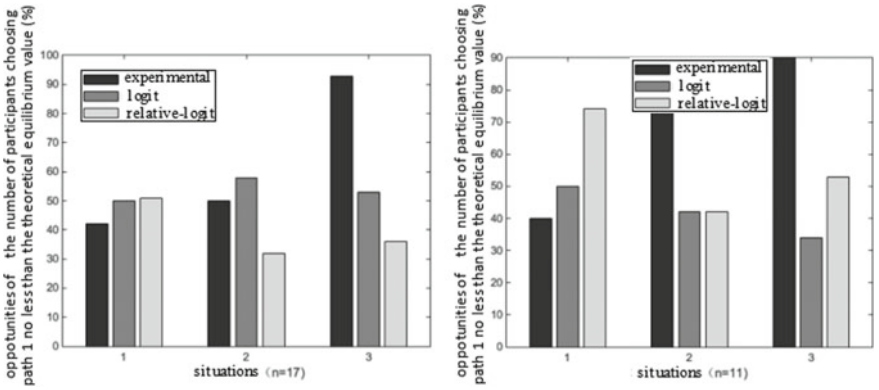


Fig. 11 Comparison of opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value (Opp)

difference become more obvious is as follow: participants generally indicated that they preferred route 1 more because it seemed obvious that route 1 was shorter, so even if more participants on route 1 than route 2, the probability of choosing route 1 to achieve shorter travel times was greater than choosing route 2.

Therefore, let's assume that the influential factors of perceived travel time not only include the historical experience and traffic information, but also the property of the routes itself (specifically, the length of each route in this paper). And modify Formula (9) to be: the expected travel time on day $i + 1$ expressed to be the weighted sum of the actual travel time on day i, the expected travel time on day i and the influence of route length, which is shown as follow:

$$\mu_{\mathrm{k}}^{(\mathrm{n,i}+1)} = \beta_{\mathrm{t}}\tilde{\mathrm{t}}_{\mathrm{k}}^{(\mathrm{n,i})} + \beta_{\mu}\mu_{\mathrm{k}}^{(\mathrm{n,i})} + \beta_{L}\gamma\tilde{L}_{\mathrm{k}} \tag{11}$$

From above, $\mu_{\mathrm{k}}^{(\mathrm{n,i}+1)}$ is the expected value of perceived travel time of traveler n on day i + 1, $\tilde{\mathrm{t}}_{\mathrm{k}}^{(\mathrm{n,i})}$ is the corresponding perceived value of actual travel time on route k on day i, $\mu_{\mathrm{k}}^{(\mathrm{n,i})}$ is the expected value of perceived travel time of traveler n on day i, $\beta_{\mathrm{t}} + \beta_{\mu} + \beta_{L} = 1$, $\beta_{\mathrm{t}}, \beta_{\mu}, \beta_{L} \in [0, 1]$, characterizing the trade-off of participant agent n among 3 influential factors, $\tilde{L}_{\mathrm{k}}$ is the corresponding perceived value of the length of route k, γ is the coefficient, characterizing the perceived influence of route length on travel time.

4.2 The Relationship Between Psychological Perception and Physical Quantity

According to Stevens's law [27], psychological perception value S is a power function of physical quantity I, which means the value perceived is directly in proportional to the power of stimulus quantity:

$$S = \mathrm{b}I^{\mathrm{a}} \tag{12}$$

From above, S is the perception value, b is a constant coefficient, a is the power index determined by the perceiving channel and the intensity of stimulus.

Thus, the traveler's perception of route length and actual travel time is:

$$\tilde{L}_{\mathrm{k}} = b_{\mathrm{k}}(L_{\mathrm{k}})^{\mathrm{a_k}} \tag{13}$$

$$\tilde{\mathrm{t}}_{\mathrm{k}} = b_{\mathrm{t}}(\mathrm{t_k})^{\mathrm{a_t}} \tag{14}$$

From above, $\tilde{L}_k$ is the corresponding perceived value of the length of route k, L_k is the length of route k, b_{k}, b_{t} is the corresponding constant coefficient, $\mathrm{a_k}$, $\mathrm{a_t}$ is the corresponding the power index.

4.3 Simulation and Analysis

Replace the original decision mechanism in the multi-agent model with modified decision mechanism described in Sects. 3.1 and 3.2, set $\lambda = 0.5$, $\beta_t = 0.2$, $\beta_\mu = 0.3$, $\beta_L = 0.5$, $\gamma = 0.02$, $b_{\mathrm{t}} = 1.2$, $\mathrm{a_t} = 1$, $b_{\mathrm{k}} = 0.8$, $\mathrm{a_k} = 2$. Take situation 2 as an example for illustration, set the number of participant agents to be $n = 17$, the initial number of participants on route 1 is 11, the number of participants on route 1 and 2 as time pass by is shown in Fig. 12.

The average number of participants on route 1 (Ave) and the opportunities of the participants on route 1 no less than the theoretical equilibrium value (Opp) with the criterion of cognition being modified and the probability calculated through logit model and relative logit model was shown in Tables 6 and 7.

The comparison of average number of participants on route 1 when $n = 17$, $n = 11$, with the probability calculated through logit model, relative-logit model before and after modified the criterion of cognition and experimental results was shown in Fig. 13. The comparison of opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value when $n = 17, n = 11$, with the probability calculated through logit model, relative-logit model before and after modified the criterion of cognition and experimental results was shown in Fig. 14.

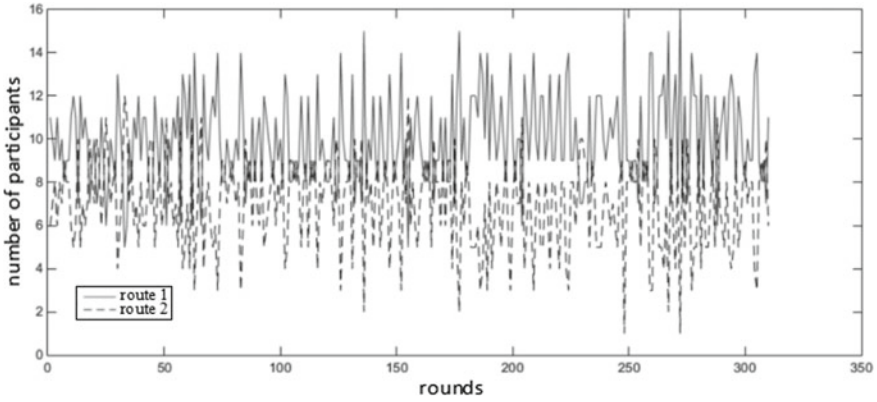


Fig. 12 The changing of the number of participants choosing path 1 under multi-agent simulations (after modified)

Table 6 Average value of the number of participants choosing path 1 (Ave) and Opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value (Opp) with logit model (after modified)

Number of participants on route 1	Situation 1		Situation 2		Situation 3	
	n = 17	n = 11	n = 17	n = 11	n = 17	n = 11
Ave	8.8	5.3	9.7	8.2	12.1	8.7
Opp (%)	49	50	52	68	84	92

Table 7 Average value of the number of participants choosing path 1 (Ave) and Opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value (Opp) with relative logit model (after modified)

Number of participants on route 1	Situation 1		Situation 2		Situation 3	
	n = 17	n = 11	n = 17	n = 11	n = 17	n = 11
Ave	8.5	5.1	9.4	7.8	11.8	8.9
Opp (%)	51	66	57	75	89	83

We may see that the results of simulation with the decision making criterion of cognition being modified fit with the experimental results better both on average number of participants on route 1 and the opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value, with the probabilities calculated through logit model and relative-logit model under all 3 route length situations. which means the decision-making mechanism with the traveler's criterion of cognition being modified reflects the fact that travelers tend to choose the shorter length route more when length difference exists, and this phenomenon become more significant when the difference become more obvious better.

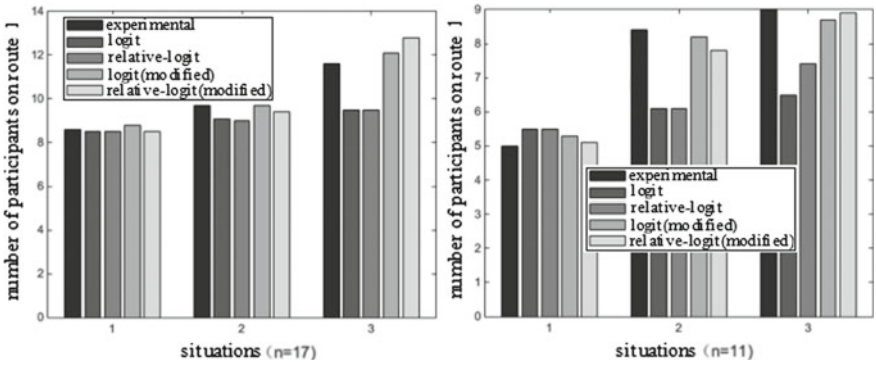


Fig. 13 Comparison of average value of the number of participants choosing path 1 (Ave)

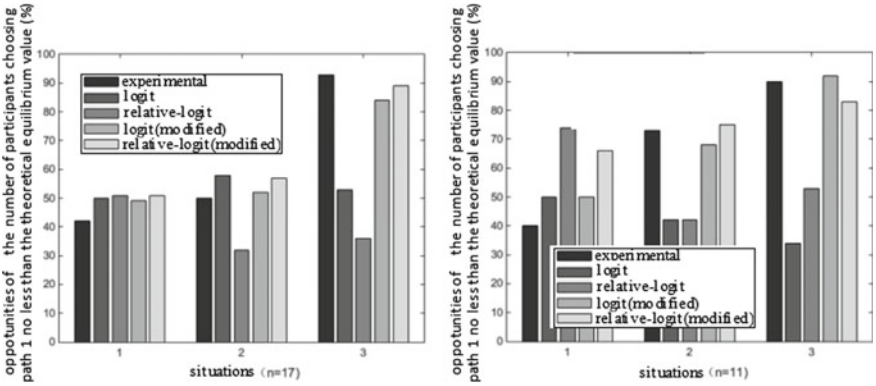


Fig. 14 Comparison of opportunities of the number of participants choosing path 1 no less than the theoretical equilibrium value (Opp)

5 Conclusions

This paper focuses on the 'equilibrium dissension' and the 'decision making mechanism' under the condition of historical experience and traffic information by conducting human-computer interaction experiments and multi-agent simulations. This study is beneficial for understanding the traveler's route choice behavior and the causes of traffic congestion.

The conclusions and the following research directions are as follows:

(1) the results of experiments in this paper support the conclusion of "no convergence to equilibrium".

(2) The results of multi-agent simulations show that the existing route choice mechanism support the conclusion of "no convergence to equilibrium" as well. Dissensions may due to exact treatment on the probability of logit discrete choice model.
(3) The decision-making mechanism proposed by existing literature which treat the influential factors of perceived travel time only include the historical experience and traffic information is not sufficient to reflect the fact that travelers tend to choose the shorter length route more when length difference exists, and this phenomenon become more significant when the difference become more obvious.
(4) The decision-making mechanism with the traveler's criterion of cognition being modified takes into consideration the property of the routes itself as one of the influential factors of perceived travel time except for the historical experience and traffic information, the after modified mechanism reflects the fact that travelers tend to choose the shorter length route more when length difference exists, and this phenomenon become more significant when the difference become more obvious better.
(5) How the 3 influential factors, the property of the routes, the historical experience and traffic information works? Are there more factors that are influential? How about the law for their functioning? These questions are the following-up research direction of this paper.

References

1. Patriksson M (1994) The traffic assignment problem-model and method. VSP, The Netherlands
2. Watling D (1999) Stability of the stochastic equilibrium assignment problem: a dynamical systems approach. Transp Res B 33(4):281–312
3. Zhou B, Xu M, Meng Q, Huang Z (2017) A day-to-day route flow evolution process towards the mixed equilibria. Transp Res Part C 82:210–228
4. Zhu W, Ma S, Tian J, Li G (2016) Nonlinear relative-proportion-based route adjustment process for day-to-day traffic dynamics: modeling, equilibrium and stability analysis. Commun Nonlinear Sci Numer Simulat 40:129–137
5. Hazelton M, Watling D (2004) Computation of equilibrium distributions of Markov traffic assignment models. Transp Sci 38(3):331–342
6. Guan HZ, Pu L (2010) A drivers' choice behavior model based on evolutionary game theory. J Beijing Univ Technol 36(8):1077–1083
7. Liu TL (2007) Multi-agent simulation on day-to-day route choice behavior. Acta Phys Sin 56(11):492–4986
8. Selten R, Chmura T, Pitz T, Kub S, Schreckenberg M (2007) Commuters route choice behaviour. Games Econ Behav 58(2):394–406
9. Avineri E, Bovy PHL (2008) Identification of parameters for a prospect theory model for travel choice analysis. Transp Res Rec J Transp Res Board 2082(2082):141–147
10. Kim H, Lim Y (2012) A Day-to-day route choice model based on drivers' past experience. KSCE J Civ Eng 16(7):1267–1279
11. Chorus CG (2012) Regret theory-based route choices and traffic equilibria. Transportmetrica 8(4):291–305

12. Di X, Liu H X, Zhu S et al (2016) Indifference bands for boundedly rational route switching. Transportation, 1–26
13. Zhao C, Huang H (2016) Experiment of boundedly rational route choice behavior and the model under satisfying rule. Transp Res Part C Emerg Technol 10(68):22–37
14. Wei F, Ma S, Jia N (2014) A day-to-day route choice model based on reinforcement learning. mathematical problems in engineering (2014-9-30) (3)
15. Meneguzzer C, Olivieri A (2013) Day-to-day traffic dynamics: laboratory-like experiment on route choice and route switching in a simple network with limited feedback information. Procedia Soc Behav Sci 87:44–59
16. Maio MLD, Vitetta A, Watling D (2013) Influence of experience on users' behaviour: a day-to-day model for route choice updating. Procedia Soc Behav Sci 87:60–74
17. Yang L, Wang H, Zhu S (2016) Summary and development trend of traffic equilibrium research. In: International conference on green intelligent transportation system and safety. Springer, Singapore, 537–548
18. Tisue S, Wilensky U (2004) NetLogo: a simple environment for modeling complexity. In: International conference on complex systems, 16–21
19. Damaceanu RC (2011) An agent-based computational study of wealth distribution in function of technological progress using Netlogo. Am J Econ 1(1):15–20
20. Fioretti G (2013) Romulus-Catalin Damaceanu: agent-based computational economics using netlogo. J Evol Econ 23(3):689–692
21. Huang H, Liu T, Yang H (2008) Modeling the evolutions of day-to-day route choice and year-to-year ATIS adoption with stochastic user equilibrium. J Adv Transp 42(2):111–127
22. Cantarella GE, Cascetta E (1995) Dynamic processes and equilibrium in transportation networks: towards a unifying theory. Transp Sci 29(4):305–329
23. Yang H, Kitamura R, Jovanis PP et al (1993) Exploration of route choice behavior with advanced traveler information using neural network concepts. Transportation 20(2):199–223
24. Xin-Jun LAI, Zhi YU, Jun LI (2012) Derivation, Implementation and Examination of Logit Route Choice Model with Relative Impedance. J Transp Syst Eng 12(2):85–90
25. Tawfik AM, Rakha HA (2014) Can we model driver perceptions? An in-situ experiment in real-world conditions. Int J Transp Sci Technol 3(2):149–165
26. Kim H, Oh JS, Jayakrishnan R (2009) Effects of user equilibrium assumptions on network traffic pattern. KSCE J Civ Eng 13(2):117–127
27. Guo X (2007) Experimental psychology

Intersection Traffic Signal Optimization Considering Lane-Changing Behavior Caused Nearby Bus Bay Stop Upstream

Rui Li, Xin Xue, Linchao Li, Changjiang Zheng and Jinxing Shen

Abstract Transit-oriented traveling mode is recognized as one of the most effective strategies for improving traveling service level and decreasing travel times, stops and delay. However, the transit system will also bring some new bottlenecks. The influence caused by bus bay stop upstream should be considered for traffic signal optimization. This paper proposes a traffic signal optimization model for intersection and bus bay stop upstream unit, which targets for minimizing traveling time and considering the effects caused by bus bay stop upstream. First, the equivalent volume method is addressed for quantified the influences generating by transits' lane-changing. Then, the influences are taken into account when solving the traffic signal optimization algorithm. Finally, the proposed optimization model is evaluated using a VISSIM model calibrated with field traffic volume and traffic signal data of the intersection of Qingliangmen Boulevard and Nenjiang Road in Nanjing, China. The evaluation results illustrate that analytical and simulation calculation results have similar performance in reducing traveling delay. Therefore, the proposed traffic signal optimization method performs well at the intersection with bus bay stops upstream.

Keywords Transport engineering · Intersection signal optimization · Near-side bus bay stop · Transits' lane-changing behavior

R. Li (✉) · X. Xue · C. Zheng · J. Shen
College of Civil and Transportation Engineering, Hohai University, Xi Kang Road 1, 210098 Nanjing, China
e-mail: lirui2012@hhu.edu.cn

Jiangsu Province Research Center for Traffic and Infrastructure Engineering Technology, Xi Kang Road 1, 210098 Nanjing, China

L. Li
School of Transportation, Southeast University, Southeast University Road 2, 211189 Nanjing, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_14

1 Introduction

With the increasing traveling demand, traditional vehicle-oriented traveling mode has been one of the most challenging for traffic congestion in China. Transit-oriented traveling mode is recognized as one of the most effective strategies for decreasing the level of traffic congestion and improving traveling efficiency. However, the transit traveling mode will also result in some new problems (especially nearby bus bay stop), transits' lane-changing behavior will decrease general vehicles' traffic efficiency. Previous experience has shown that bus stop upstream will significantly affect traffic characteristics of intersections, and the influence should be considered for traffic signal optimization.

Traffic signal phasing optimization for urban intersections has been extensively investigated in the past decades. Passengers' average delay [1], vehicles' average delay, passengers' total delay, traveling punctuality [2], and regularity [3] are analyzed as the target for optimizing. Furthermore, the impacts of traffic flow characteristics [4], geometry configurations [5, 6], and lane-changing behaviors [7] have also been used. Extending the celebrated kinematic wave theory to incorporate lane-changing is a natural idea and has been attempted since the 1970s [8–12]. Then, the famous fundamental models for analyzing lane-changing behavior's impact on surrounding vehicles was initially proposed by Newell [13] and later formalized by Daganzo [14], and qualitative conjectures [15], empirical evidence [16], and quantitative understanding [7] model has been developed. While serving passengers at a bus stop, buses can interact on the bus lanes that limit their discharge flows. Geometry configurations [17], traffic flow patterns [6], and other characteristics [18] near bus stop are analyzed and simulated [19] for quantifying the capacity drop near the scope of bus service stop. However, the interaction caused by the lane-changing behaviors at the segment between approaches and bus bay stop upstream will strongly influence the traffic flow pattern, and the influence to traffic signal optimization strategy should be paid more attention and needs further development.

In this research, we attempt to develop a traffic signal optimization method considering the lane-changing behavior influences caused nearby bus bay. The proposed optimization framework is formulated concerning the intersection and near-side bus bay as depicted in Fig. 1 (defined as the I-SU unit). For the I-SU unit, original traffic flow characteristics are usually changed during the processes of transits' lane-changing at the bus stop. According to the analysis and quantification of the influences caused by transits' lane-changing behaviors for traveling time increase and capacity drop, the signal phasing optimization method is developed for I-SU unit.

The rest of this paper is organized as follow. In Sect. 2, traffic signal optimization functions for intersection is proposed, effects caused by near-side bus bay are analyzed and quantified, and traffic signal phasing optimization method is developed. Section 3 is the field-data based analysis, calibration, and evaluation of the proposed optimization scenarios. The performance is also analyzed and simulated by VISSIM-based simulation platform. In the end, conclusions and recommendations are included.

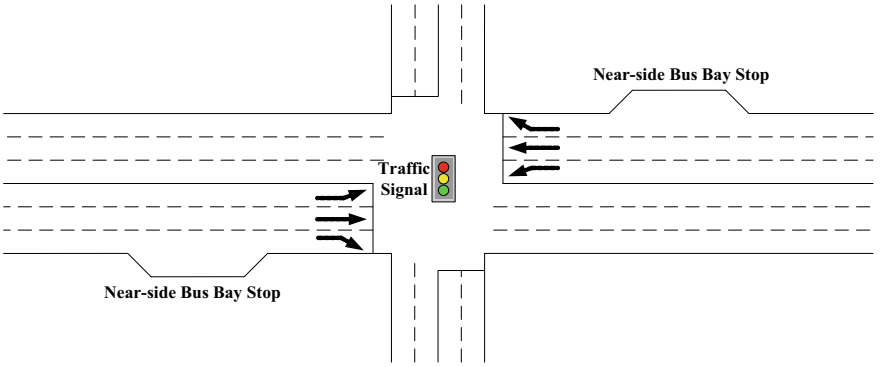


Fig. 1 Visualization of the intersection with bus bay stop upstream

2 Methodology

The I-SU unit will obviously change the traffic volume patterns on the approach lane near bus stops. The processes of transits' lane-changing for entering and exiting bus bay (shown in Fig. 2) will directly affect original traffic flow characteristics and increase vehicles' traveling time. The research in this paper develops an improved traffic signal optimization model for intersection with bus bay stop upstream. For the I-SU unit signal optimization model, minimize total traveling delay (of general vehicles and transit vehicles) is proposed as the optimization target, and the impacts caused by lane-changing behavior nearby bus bay upstream for traffic volume patterns are quantified.

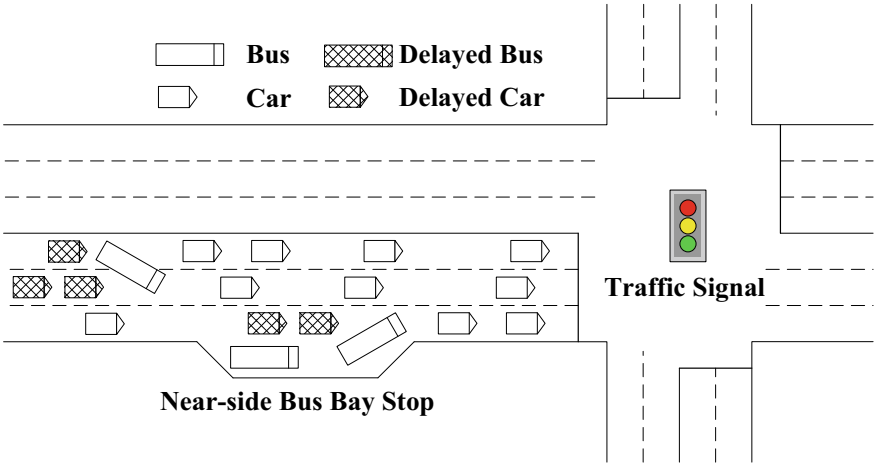


Fig. 2 Illustrative extra traveling time generation caused by lane-changing nearby bus bay stop

2.1 Intersection Traffic Signal Optimization Target

In this research, a traffic signal optimization model considering lane-changing behavior nearby bus bay upstream is proposed. Transit service bay resides upstream of its nearby intersection may become a bottleneck that constrains vehicle flows near the intersection. Minimum vehicles total delay (D) of the intersection is the objective function for intersection traffic signal optimization.

$$\begin{aligned} &\min D = \sum_{h\in H}\sum_{j\in J} D_I^{hj}(G, T, V) \\ &s.t.\ \max\tfrac{x_{hj}}{\lambda_{hj}} \le 0.9 \quad h \in H,\ \ j \in J \end{aligned} \tag{1}$$

$D_I^{hj}(G, T, V)$ denotes the passenger delay at intersections (for the phase j of cycle h) under the condition of geometry configurations (G), signal phasing (T), and traffic flow pattern (V), and the duration of green time is the optimization variable. H is the set of all cycles of the intersection at a particular time, J is the set of all phases of one cycle of the intersection at a specific time, x_{hj} and λ_{hj} are the degree of saturation and green time ratio of general vehicles respectively for phase j of cycle h.

The total delay at intersection consists of the traveling delay of general vehicles and transit vehicles. The total traveling delay for phase j of cycle h (D_I^{hj}) is addressed as the following Equation.

$$D_I^{hj} = d_{Ig}^{hj} \times q_{Ig}^{hj} + d_{It}^{hj} \times q_{It}^{hj} \tag{2}$$

where d_{Ig}^{hj} and d_{It}^{hj} are the average delay (for lane group) of general vehicles and transit vehicles for phase j of cycle h. q_{Ig}^{hj} and q_{It}^{hj} are the average arrival rate (for lane group) of general vehicles and transit vehicles for phase j of cycle h.

The average vehicle-based delay for lane group can be calculated based on the formula cited from Highway Capacity Manual [22]. The average delay of vehicles for lane group at one signalized intersection is illustrated in Eq. (3).

$$d_I = \frac{0.5C(1-\lambda)^2}{1-[\min(1, X)\lambda]} + 900T\left[(X-1) + \sqrt{(X-1)^2 + \frac{8klX}{cT}}\right] \tag{3}$$

$$X = \frac{q}{c} \tag{4}$$

where C is the cycle length, λ is the green time ratio for lane group, X is the degree of saturation for lane group, T is the analysis duration, k is the incremental delay factor that is dependent on controller settings, l is the upstream filtering adjustment factor, c is the lane group capacity, and q is the arrival rate of vehicles for lane group. The degree of saturation and capacity for lane group are recalibrated in the following research considering the effects of lane-changing behavior caused near-side bus bay stop.

2.2 *Effects Analysis Caused by Lane-Changing Behavior Near-Side Bus Bay Stop*

Traffic flow characteristics usually change around the area of transits' lane-changing near the bus stop. For transits' lane-changing behavior process, extra traveling times of following vehicles on the subject and target lane will generate, and the effects of these extra traveling times will be analyzed using equivalent volume method.

2.2.1 Transits' Lane-Changing Behavior Influence Analysis

According to analyze the extra traveling time of subject and target lane caused by lane-changing behavior, the following vehicles with significant influence are determined. Then, the equivalent-flow analysis method is used for quantifying the influence of traffic flow characteristics caused by transit's lane-changing behavior.

(1) Extra traveling time analysis

Figure 3 shows the transit vehicles' ($B_I(S)$ and $B_O(S)$) lane-changing process for entering into bus stop and driving out from the bay. Transits' lane-changing behaviors will significantly increase the traveling time of following vehicles moving on the subject and target lane.

For the lane-changing process of buses' entering into bus bay, the following vehicles in the subject lane are labeled as $F_{IC}(S1), F_{IC}(S2), \ldots, F_{IC}(Sm)$, and the follow vehicles moving in the subject lane are labeled as $F_{IC}(T1), F_{IC}(T2), \ldots, F_{IC}(Tn)$. The last affected vehicle of the following fleet on the subject lane and target lane are m and n, respectively. For the lane-changing process of transit vehicles' driving out of the bus bay, the following vehicles in the subject lane are labeled as $F_{OC}(T1), F_{OC}(T2), \ldots, F_{OC}(Tm)$. Taking the lane-changing analysis of buses' entering into bus bay as an example.

Blocked vehicles' extra traveling time for the process of buses' entering into bus bay is determined by analyzing their trajectory. The extra traveling time is defined as the increased time, which is calculated by comparing with the actual and original

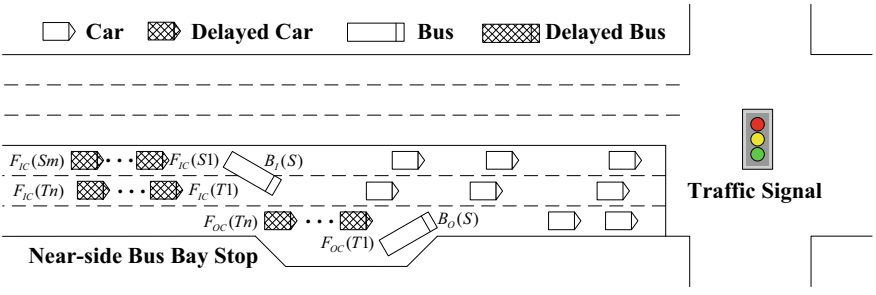


Fig. 3 Illustrative lane-changing behavior of transit

traveling time moving from lane-changing generation spot to approach stop bar. $\Delta t_{B_I(S)}$, $\Delta t_{F_{IC}(S1)}$, and $\Delta t_{F_{IC}(Sm)}$ are the extra traveling time of the lane-changing bus, the first and last following vehicle moving in the subject approach. The extra traveling time for transit and following vehicles are shown in Eqs. (5) and (6).

$$\Delta t_{B_I(S)} = T^D_{B_I(S)} - \hat{T}^D_{B_I(S)} \tag{5}$$

$$\begin{cases} \Delta t_{F_{IC}(S1)} = T^D_{F_{IC}(S1)} - \hat{T}^D_{F_{IC}(S1)} \\ \vdots \\ \Delta t_{F_{IC}(Sm)} = T^D_{F_{IC}(Sm)} - \hat{T}^D_{F_{IC}(Sm)} \end{cases} \tag{6}$$

where $T^D_{B_I(S)}, T^D_{F_{IC}(S1)}$, and $T^D_{F_{IC}(Sm)}$ are the actual arrival time of the lane-changing bus, the leader following vehicle, and the last following vehicle on the subject approach. $\hat{T}^D_{B_I(S)}$, $\hat{T}^D_{F_{IC}(S1)}$, and $\hat{T}^D_{F_{IC}(Sm)}$ are the original arrival time of the lane-changing transit vehicle, the leader following vehicle, and the last following vehicle on the subject approach.

Be similar to the methodology for following vehicles' extra traveling time calculation, the following vehicles' extra traveling time caused by lane-changing on target lane is illustrated in Fig. 3.

$$\begin{cases} \Delta t_{F_{IC}(T1)} = T^D_{F_{IC}(T1)} - \hat{T}^D_{F_{IC}(T1)} \\ \vdots \\ \Delta t_{F_{IC}(Tn)} = T^D_{F_{IC}(Tn)} - \hat{T}^D_{F_{IC}(Tn)} \end{cases} \tag{7}$$

where $T^D_{F_{IC}(T1)}$ and $T^D_{F_{IC}(Tn)}$ are the actual arrival time of the leader following vehicle and the last following vehicle on the target lane, while $\hat{T}^D_{F_{IC}(Tn)}$ and $\hat{T}^D_{F_{IC}(Tn)}$ are the original arrival time of the leader following vehicle and the last following vehicle on the target lane.

Be similar to the methodology for entering into bus bay vehicles' extra traveling time calculation, the extra traveling time caused for the process of driving out of the bus stop for transit, following vehicles on subject lane and target lane are shown in Eqs. (8), (9) and (10), respectively.

$$\Delta t_{B_O(S)} = T^D_{B_O(S)} - \hat{T}^D_{B_O(S)} \tag{8}$$

$$\begin{cases} \Delta t_{F_{OC}(S1)} = T^D_{F_{OC}(S1)} - \hat{T}^D_{F_{OC}(S1)} \\ \vdots \\ \Delta t_{F_{OC}(Sm)} = T^D_{F_{OC}(Sm)} - \hat{T}^D_{F_{OC}(Sm)} \end{cases} \tag{9}$$

$$\begin{cases} \Delta t_{F_{OC}(T1)} = T^D_{F_{OC}(T1)} - \hat{T}^D_{F_{OC}(T1)} \\ \vdots \\ \Delta t_{F_{OC}(Tn)} = T^D_{F_{OC}(Tn)} - \hat{T}^D_{F_{OC}(Tn)} \end{cases} \tag{10}$$

(3) Equivalent volume calculation

The effects of these extra traveling times caused by lane-changing behavior of transit vehicles can be analyzed with equivalent volume method. Equivalent volume (Q) includes basic volume (Q^b) and additional volume (Q^a). Basic volume is the real traffic volume. The additional volume is the volume considering increased traveling time, which includes the additional volume on the subject lane and target lane for the transit's lane-changing process.

$$Q = Q^b + Q^a \tag{11}$$

$$\begin{cases} Q^a = \sum\limits_{q^a_I \in Q^a_I} q^a_I + \sum\limits_{q^a_O \in Q^a_O} q^a_O \\ q^a_I = \frac{\Delta t_{F_{IC}(S1)} + \Delta t_{F_{IC}(S2)} + \cdots + \Delta t_{F_{IC}(Sm)}}{m \cdot \bar{h}_S} + \frac{\Delta t_{F_{IC}(T1)} + \Delta t_{F_{IC}(T2)} + \cdots + \Delta t_{F_{IC}(Tn)}}{n \cdot \bar{h}_T} \\ q^a_O = \frac{\Delta t_{F_{OC}(S1)} + \Delta t_{F_{OC}(S2)} + \cdots + \Delta t_{F_{OC}(Sm)}}{m \cdot \bar{h}_S} + \frac{\Delta t_{F_{OC}(T1)} + \Delta t_{F_{OC}(T2)} + \cdots + \Delta t_{F_{OC}(Tn)}}{n \cdot \bar{h}_T} \end{cases} \tag{12}$$

where q^a_I and q^a_O denote the additional volume on the lanes (subject lane and target lane) for entering into bus stop and driving out of the bus bay, $\bar{h}_S$ and $\bar{h}_T$ denote the headway of the following fleet on the subject lane and target lane.

2.3 Traffic Signal Phasing Optimization

The framework of intersection traffic signal phasing optimization is shown in Fig. 4. Firstly, initial signal phasing plan for intersection is proposed, and the total traveling delay considering lane-changing behavior caused nearby bus bay of intersection for the initial plan is calculated. Then, the improved signal phasing plan for intersection is formed, and the traveling delay for the improved plan is determined. If the vehicle traveling delay of the initial plan is more than that of the improved plan, the initial plan will be replaced by the improved plan, and signal plan optimization will continue. Otherwise, the improved signal phasing plan is the optimization plan.

In this paper, the improved signal plan will be proposed by adjusting the green time for signal phases [20]. The green time for the improved signal plan can be calculated by Eq. (13).

$$g = \begin{cases} g_k + 1 & k = a \\ g_k\left(1 - \frac{1}{C - L - g_a}\right) & k \neq a \end{cases} \tag{13}$$

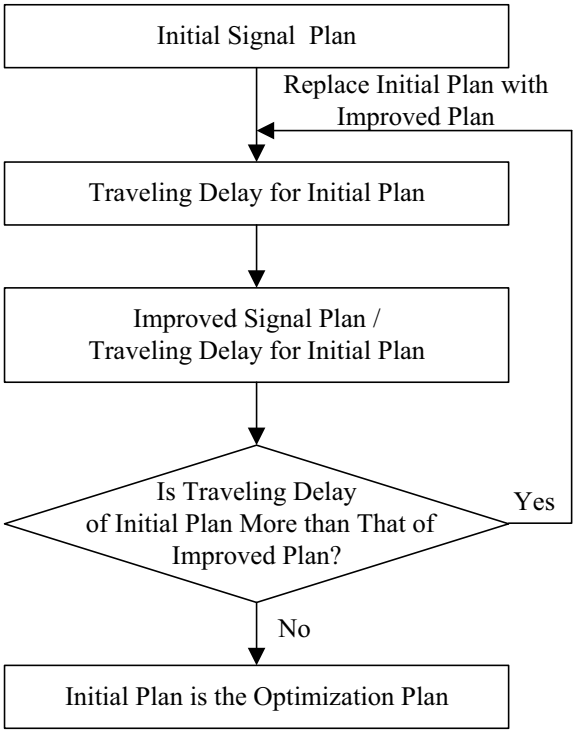


Fig. 4 Framework of signal phasing optimization

where a is the major improved signal phase, L is the lost time in each cycle, g_k and g_a are the green times of the initial phasing plan in phase k and a, k is the set of signal phase.

3 Case Study

3.1 Experiment Design and Data Description

In this study, field data collected from the intersection of Qingliangmen Boulevard and Nenjiang Road in Nanjing, China are used in the numerical experiments for the proposed traffic signal optimization model. Figure 5 illustrates the geometric layout and original signal phase plan for the intersection of Qingliangmen Boulevard and Nenjiang Road. The main direction locates on Qingliangmen Boulevard, and near-side bus bay stop (Zhanjiang bus stop) of Qingliangmen Boulevard for this intersection serve for seven bus routes (Bus Routes 15, 23, 60, 91, 133, 303, and 317). These seven transit routes have an average headway of 5–10 min during the morning rush hour.

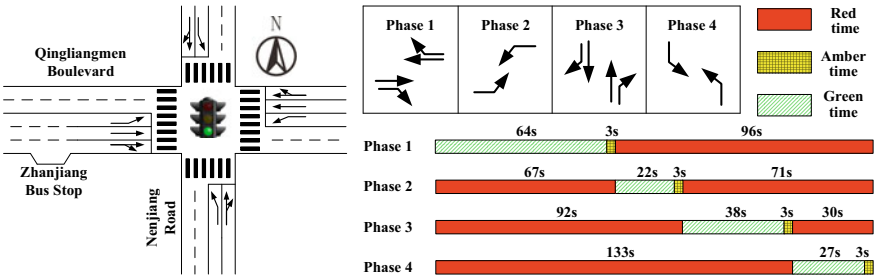


Fig. 5 Layouts and phasing for the experimental intersection

Table 1 Traffic volumes in the morning rush hour for the experimental intersection

Approach	Vehicle type	March 7nd			March 9rd		
		LT	TH	RT	LT	TH	RT
Eastbound	General vehicle	52	883	58	36	895	63
	Transit vehicle	0	51	0	0	64	0
Westbound	General vehicle	90	432	134	99	508	119
	Transit vehicle	4	60	13	4	67	13
Southbound	General vehicle	60	58	113	74	56	121
	Transit vehicle	0	0	3	0	0	6
Northbound	General vehicle	140	96	66	185	90	75
	Transit vehicle	6	0	0	6	0	0

Traffic volumes in the morning rush hour (7:30–8:30 am) are shown in Table 1. The flow data of general and transit vehicles are collected on March 7nd and 9th, 2017.

Lots of preliminary experiments were performed to develop the basic capacity of lanes which is 1700 vehicles per lane per hour [21], and the transit lane basic capacity is 850 buses per lane per hour [22]. During the morning rush hours of two observation days, there are 79 and 62 lane-changing behaviors for entering into bus stop and driving out of bus bay occurred in eastbound approach.

3.2 Experimental Results and Analysis

According to analyze the increased traveling time caused by lane-changing behaviors nearby the bus stop, the effects caused by laying out near-side bus service stop are determined. The additional equivalent volumes of eastbound approaches is 147 pcu, and the optimization signal phasing plan for this observed intersection is proposed in Fig. 6.

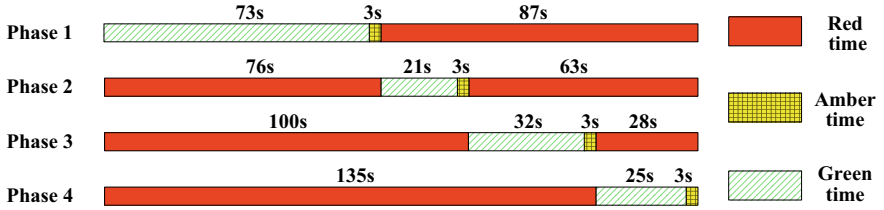


Fig. 6 Optimization signal phasing plan for the experimental intersection

The VISSIM-based simulated environment of this study is according to the study conducted by Li et al. [20]. Original and Optimal phasing plans for the intersection of Qingliangmen Boulevard and Nenjiang Road are simulated. Total delay for this intersection with bus stop upstream are determined using analytical method (proposed in this paper) and simulation method (VISSIM-based), the results demonstrate the significant reduction of nearly 5.6% in delay reduction for the test intersection, analytical and simulation calculation results have similar performance in reducing traveling delay. Therefore, the proposed traffic signal optimization method performs well at the intersection with the near-side bus bay stop.

4 Conclusion

This paper proposes a traffic signal optimization framework for I-SU unit, which considers the effects caused by the near-side bus stop. The processes of transits' lane-changing at bus stops are focused on for analyzing the changes of traffic flow characteristics caused by bus bay stop upstream. The equivalent-flow analysis method is used for quantifying the influence of traffic flow characteristics caused by transit's lane-changing behavior. The framework of traffic signal phasing optimization is illustrated. The evaluation results illustrate that analytical and simulation calculation results have similar performance in reducing traveling delay. Therefore, the proposed traffic signal optimization method performs well at the intersection with the near-side bus bay stop. Future work includes more extensive numerical experiments and field tests to assess the effectiveness of the proposed signal optimization method for I-SU unit.

Acknowledgements This research was supported by the Natural Science Foundation of Jiangsu Province (Grant No. BK20181307), Postdoctoral Science Foundation of China (Grant No. 2018M630505) and National Natural Science Foundation of China (Grant No. 51508161).

Conflict of Interests
The authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. Christofa E, Papamichail I, Skabardonis A (2013) Person-based traffic responsive signal control optimization. IEEE Trans Intell Transp Syst 14(3):1278–1289
2. Furth P, Muller T (2000) Conditional bus priority at signalized intersections: better service quality with less traffic disruption. Transp Res Rec 1731:23–30
3. Hounsell N, Shrestha B (2012) A new approach for co-operative bus priority at traffic signals. IEEE Trans Intell Transp Syst 13(1):6–14
4. Zeng X, Zhang Y, Balke KN, Yin K (2014) A real-time transit signal priority control model considering stochastic bus arrival time. IEEE Trans Intell Transp Syst 15:1657–1666
5. Gu W, Li Y, Cassidy MJ, Griswold JB (2011) On the capacity of isolated, curbside bus stops. Transp Res Part B Methodol 45(4):714–723
6. Yang X, Gao Z, Zhao X, Si B (2009) Road CAPACITY AT BUS STOPS WITH MIXED TRAFFIC FLOW IN China. Transp Res Rec 2111:18–23
7. Laval JA, Daganzo CF (2006) Lane-changing in traffic streams. Transp Res Part B Methodol 40(3):251–264
8. Munjal PK, Pipes LA (1971) Propagation of on-ramp density perturbations on unidirectional two- and three-lane freeways. Transp Res Part B Methodol 5(4):241–255
9. Munjal PK, Hsu Y, Lawrence RL (1971) Analysis and validation of lane-drop effects on multi-lane freeways. Transp Res Part B Methodol 5(4):257–266
10. Daganzo CF (1997) A continuum theory of traffic dynamics for freeways with special lanes. Transp Res Part B Methodol 31(2):83–102
11. Daganzo CF (2002) A behavioral theory of multi-lane traffic flow part I: long homogeneous freeway sections. Transp Res Part B Methodol 36(2):131–158
12. Jin W (2010) A kinematic wave theory of lane-changing traffic flow. Transp Res Part B Methodol 44(8):1001–1021
13. Newell GF (1982) Applications of queueing theory, 2nd edn. Chapman & Hall, New York, U.S.
14. Daganzo CF (1995) The cell transmission model, part II: network traffic. Transp Res Part B Methodol 29(2):79–93
15. Chang GL, Kao YM (1991) An empirical investigation of macroscopic lane-changing characteristics on uncongested multilane freeways. Transp Res Part A Gen 25(6):375–389
16. Cassidy MJ, Rudjanakanoknad J (2005) Increasing capacity of an isolated merge by metering its on-ramp. Transp Res Part B Methodol 39(7):896–913
17. Gu W, Gayah VV, Cassidy MJ, Saade N (2014) On the impacts of bus stops near signalized intersections: models of car and bus delays. Transp Res Part B Methodol 68(2):123–140
18. Fernández R (2010) Modeling public transport stops by microscopic simulation. Transp Res Part C Emerg Technol 18(6):856–868
19. Koshy RZ, Arasan VT (2005) Influence of bus stops on flow characteristics of mixed traffic. J Transp Eng 131(8):640–643
20. Li R, Jin PJ, Ran B (2016) Bi-objective optimization and evaluation for transit signal priority strategies at bus stop-to-stop segment. Math Probl Eng 2016:640–643
21. Highway Capacity Manual. Transportation Research Board of the National Academies, Washington, D.C. (2010)
22. Transit Cooperative Research Program (TCRP) Report 165: Transit capacity and quality of service manual (third edition). Transportation Research Board of the National Academies, Washington, D.C. (2013)

Resilience Analysis for Comprehensive Transportation Network

Shuyun Niu, Ji-sheng Zhang, Fan Zhang and Jian Gao

Abstract Based on resilience theory, considering the security capability, structure capability, operation capability, emergency capability and management level of comprehensive transportation network, from the aspects of security capability, management and service capability of network, the resilience evaluation metrics of comprehensive transportation network is constructed. Meanwhile, the computing method for every index is proposed. Based on the characteristic of qualitative and quantitative combination of resilience metrics, the resilience evaluation method of comprehensive transportation network, which combines the AHP (Analytic Hierarchy Process) and TOPSIS (Technique Order Preference Ideal Solution) is proposed. The result of resilience can provide technical support for improving the reliability and preventing and control the risk of the comprehensive transportation network.

Keywords Resilience of network · Comprehensive transportation network · Evaluation metrics · Security capability of network · Operation capacity of network

1 Introduction

Recent years, the serious accident and emergency occurs frequently in China. Such as the frozen disaster of 2008 in South of China, heavy snow of 2010 in North of China, which often cause regional transportation channel broken, even wider scope traffic disruption. The purpose of this study is to evaluate the resilience of comprehensive transportation network, to provide technical support for prevention and control pre-emergency. It's the foundation of improving the reliability of comprehensive transportation network.

The word "resilience" comes from Latin "resilire", which means rebound, resilience [1]. The conception of resilience is involved in many fields, such as engineering, psychology, society, ecology, commerce and economics etc. [2–4].

S. Niu · J. Zhang · F. Zhang · J. Gao (✉)
National ITS Research Center, Research Institute of Highway Ministry of Transport, Beijing 100088, China
e-mail: gaojian@itsc.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_15

In the field of traffic engineering, there are also lots of research and definition about resilience performed by scholars [5–8]. Some of them, like Heaslip, think that resilience is the ability of a system to maintain its service level or to recover by itself to certain level of service within certain time framework.

Murray-Tuite proposed 10 latitudes for traffic system description [5]. They are redundancy, diversity, high active, self-organization, intensity, cooperation, adaptability, motility, security and quick restorability. Some scholars, such as Mostashari defined two kinds of resilience evaluation indexes, proposed the model framework of resilience evaluation for regional road network [9].

Up to now, the resilience studies mainly focus on the model framework of regional road network and system of evaluation index. For example, Heaslip and Louisell proposed the conception of resilience cycle which includes four stages of normal, descending, self-recovering and recovering [10]. Some researchers structure the resilience evaluation metrics from the view of vulnerability and adaptability. Some others evaluate the resilience from two aspects of travel time disturbance and environment resilience. Based on Ref. [10], Serulle and Heaslip, expanded and refined the concepts on measurement of transportation network resiliency at the pre-event level [11]. The methodology included four levels, mainly was to refine the index from two aspects of base resilience and network management.

In summary, the references mainly were from the aspect of anti-interference ability of network, and the concerned objects are network itself and users. However, under the condition of emergency, network management and service ability are also the key factor for network capability recovering. The transportation networks, which have high management and service ability, will reduce much of the time of network self-recovering and recovering stages. Therefore, based on the research of the formers, this article combines the security capability, operation capability, management and service capability of transportation network infrastructure with resilience evaluation of comprehensive transportation network.

2 The Evaluation Metrics of Resilience Construction

In order to make the evaluation results objectively and accurately reflect the resilience of the comprehensive transportation network, the systematic, fairness, practicability principles are followed in the process of evaluation index selection.

2.1 The Resilience Evaluation Metrics of Comprehensive Transportation Network

Based on above-mentioned selection principles of evaluation index, resilience evaluation metrics of comprehensive transportation network is proposed. The system

includes two aspects of security capability, management and service capability of network. Network security capability index includes structure, operation and infrastructure security capability. Among them, the infrastructure security capability is lever variable. Network structure capability includes average shortest path length, road density, node density and multimode transfer level. Network operation capability includes load margin, path diversity, average delay, average speed decent rate and traffic constitution etc. Network management and service level index includes emergency capability and management level. Among them, emergency capability covers plan completeness, level of supplies reserve, emergency supplies reachability etc. Management level includes software/hardware matching facilities, information release and service facilities and operation monitor facilities etc.

The evaluation metrics of transportation network system is as Fig. 1.

2.2 Definition of Evaluation Index and Computing Method

2.2.1 Network Security Capability

Network security capability includes three aspects of infrastructure safety capability, structure capability and operation performance.

1. **Security capability of network infrastructure**

Using the evaluation result of existing references, make the qualitative evaluation values of network infrastructure security capability corresponding to quantitative value. From low to high, the values are 1.0, 1.2, 1.4, 1.6, 2.0.

2. **Network structure capability**

It refers to connectivity of regional traffic network, density of road planning, accessibility of network and transfer convenience between multimode transportation.

(1) Average shortest path length
It refers to the average value of the shortest path among the main OD pairs in the regional traffic network. For the same road network, the smaller this value is, the better the connectivity.
(2) Comprehensive transportation network density C_d

It refers to the equivalent lanes of every 100 km^2 multiply kilometers in the regional road network. The unit is equivalent lane km/100 km^2. The calculation formula is

$$C_d = \omega_1 \times H_d + \omega_2 \times T_d + \omega_3 \times A_d \tag{1}$$

In which, C_d is the comprehensive transportation network density, H_d is the highway network density, T_d is the railway network density, A_d is the aviation network density, and the calculation formulas are as follow.

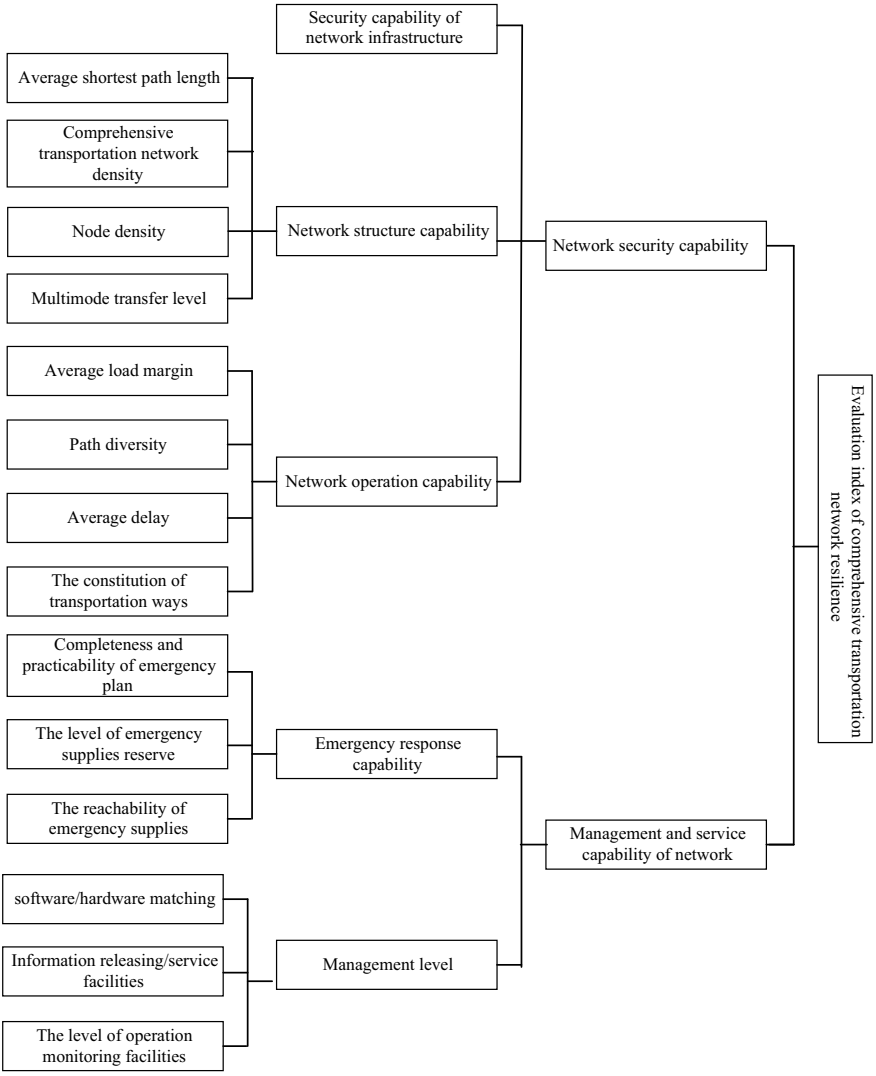


Fig. 1 Evaluation index of comprehensive transportation network resilience

$$H_d = \frac{\sum_i l_i \times n_i}{S} \times 100, \quad T_d = \frac{\sum_i l_i}{S} \times 100, \quad A_d = \frac{\sum_i l_i}{S} \times 100.$$

where, l_i is the length of road section i. n_i is the amount of lanes of road section i. S is the area.

(3) Node density N_d

It refers to the number of node of every 100 km^2 in the regional road network. The nodes mean the road crossing, interflow hub of highway, connecting node between route and railway station, airport and port etc. This index reflects the accessibility of road network. The bigger the value is, the stronger the accessibility of road network, the more alternative routes, the stronger mobility and the lower risk.

$$N_d = \frac{N}{S} \times 100 \tag{2}$$

where, N is the amount of nodes in the regional road network. S is the area.

(4) Multimode transfer level

It refers to the transfer between different travel ways. Thus, the railway station, airport, and port are considered as transfer hubs. It is influenced by the space density and density per capita of the transfer station. The space density is a ratio of amount of the transfer station to the area in the region. The density per capita is the ratio of the amount of transfer station to the amount of travelers in the region. The amount of travelers can be obtained through mobile phone signaling data.

3. **Network operation capability**

(1) Average load margin
It refers to the ratio of network capacity to the remainder capacity of network. The average load margin of road network is the average ratio of traffic capability to the remainder traffic capability. The average load margins of railway and aviation network are the average ratio of transport capability to the remainder transport capability on the key OD lines. On this base, take the transport capability of different traffic modes as the weight to make the weighted mean.

(2) Path diversity
The connecting diversity of one OD pair can be simply defined through the amount of available paths. However, the problem is if several paths share one road section, all the paths will be out of work in case of the shared road section is broken. Therefore, one factor which can embody the path overlapping need to be considered for the diversity. It is called connecting intensity which formula proposed by Di Gangi and Luongo [12], it is taken as the multiplier of path amount. When the path diversity of the road network is calculated, take the ratio of the whole road network requirement flow to OD requirement as the OD pair's weight, to calculate the average value of the path diversity of all the OD pairs as the path diversity of network level.

(3) Average delay
The delay of road section is difference between travel time of peak hour and travel time of free flow. The average value of the key section delay is taken as average delay in the road network. For railway and aviation

network, the OD delay is the difference between running time of delay and on schedule. Based on above situation, take the transport volume of different transportation ways as weight to evaluate the weight mean.

(4) The constitution of transportation ways
It refers to the constitution of transportation ways in the network of highway, railway and aviation.

2.2.2 Management and Service Capability of Network

It concludes two parts of emergency response capability and management capability.

1. **Emergency response capability**

(1) Completeness and practicability of emergency plan
The emergency plan should be as much as systematic, comprehensive and easy to practice. This article divides the completeness and practicability of emergency plan into 5 grades of best, good, ordinary, bad and worse. In the actual evaluation, the related experts should analysis the emergency plan from different aspects to fix the grade of the index.

(2) The level of emergency supplies reserve

It reflects the allocation of supplies and equipments during the implementation of transport, organization and dispatching. In all of the required supplies, if the quantity of certain supply can satisfy the requirement, the reserve ratio of this supply is 1. The reserve ratio of the supplies without reserve is 0. For general supplies, the formula of the reserve ratio is:

$p =$ reserve quantity (amount or weight)/required quantity (amount or weight).

Cite the emergency rescue recode of lager area paralytic regional traffic in the past 3–5 years; take the average required quantity of the emergency supplies as requirement in the period, to calculate the reserve ratio. The formula of the supplies reserve ratio in the regional network is,

$$P = \frac{p_1 + p_2 + \cdots + p_n}{n} \tag{3}$$

where, p_i is the reserve ratio of the ith supply.

The related relationship of the P value and the index grade shows as Table 1.

Table 1 The grade of the level of emergency supplies reserve

P value	$95\% < P \leq 100\%$	$90\% < P \leq 95\%$	$80\% < P \leq 90\%$	$60\% < P \leq 80\%$	$0\% < P \leq 60\%$
Grade	Best	Good	Ordinary	Bad	Worse

(3) The reachability of emergency supplies R_d

The reachability of emergency supplies refers to the extent to which the goods and rescue materials reach the location of the event within the time specified by the manager. Based on the length of time T required for emergency management requirements, statistics the sum of the mileage covered by each material reserve in time T, and calculates the ratio of the reachability mileage of supplies to the total mileage of the transportation network. The ratio is recorded as the reachability degree of supplies. The greater the ratio, the higher the reachability degree of supplies.

The calculation formula is

$$R_d = \frac{\sum_j \sum_i D_{ij}}{L} \tag{4}$$

where, D_{ij} is the length of the ith reachable road from the jth reserve site. L is the total mileage of the road in the regional network.

2. **Management level**

Mainly from the aspects of software/hardware matching, information releasing/service facilities and operation monitoring facilities, evaluate the management level. The more advanced the software/hardware matching is, and the more stable the technology is, the bigger of the value.

3 Evaluation of Resilience by Combining the Methods of AHP and TOPSIS

Front section structured third level evaluation metrics and some of the indexes are quantitative and some are qualitative in the resilience evaluation metrics of the comprehensive transportation network. Based on this hybrid index characteristic, choose the method of TOPSIS [13] to make comprehensive evaluation. The basic principle of TOPSIS method is to sequence the evaluated objects by measuring the distance from the objects to the optimal and worst solutions. If the object is closed to the optimal solution and far away to the worst solution, deem this situation is the best, otherwise is not. Before the resilience evaluation of the integrated network, the weight value of every level index need to be fixed. Comparing with other analysis methods, the method of AHP needs less data, is more practicable, and delivers more objective result. Therefore, the method of AHP is adopted to fix the weight of evaluation index for network resilience.

4 Resilience Analysis of Comprehensive Transportation Network of Jing-Jin-Ji Region

Accordingly, under the condition of setting emergency, evaluate the resilience of comprehensive transportation network of Jing-Jin-Ji region. Assume that there is a large area of cloud fog at the junction of Jinghu freeway and Jingjintang freeway. Around 20 km region is influenced by the cloud fog. The freeway of Jinghu, freeway of Jingjintang and G103 are blocked. The sketch shown in Fig. 2.

Since the evaluation to the management and service capability of comprehensive transportation network of Jing-Jin-Ji region needs large quantity data support, in case the data is incomplete, it is hard for the expert to judge effectively. Therefore, in this example, assume that all the evaluation indexes of the management and service capability of comprehensive transportation network in Jing-Jin-Ji region remain unchanged before and after the emergency. Namely, the value does not influence the relative result of resilience evaluation.

According to the computing formula of evaluation index and the standardization method, the results of network resilience evaluation index with and without emergency can be obtained as Table 2. The positive ideal solution was gotten under the condition of ideal and optimal status of networks and the negative ideal solution under the extreme abominable condition, like completely paralysis.

According to the weight value confirmed through AHP method, the combination weight value of the parameter is:

Fig. 2 Sketch map of the position and influence of the emergency

Table 2 The parameter value of resilience

Index	Combination weight	Without block events	With block events	Positive ideal solution	Negative ideal solution
Length of average shortest path	0.027	1	0.99	1	0
Network density	0.096	1	0.96	1	0
Node density	0.106	1	0.94	1	0
Level of multimode transfer	0.094	1	0.99	1	0
Average load margin	0.15	0.84	0.83	1	0
Path diversity	0.117	1	0.5	1	0
Average delay	0.078	0.98	0.846	1	0
Constitution of transportation ways	0.036	0.57	0.536	1	0
Completeness of plan	0.027	[0.8,1]	[0.8,1]	[0.8,1]	[0,0.2]
Level of supplies reserve	0.09	[0.8,1]	[0.8,1]	[0.8,1]	[0,0.2]
Reachability of emergency supplies	0.106	1	1	1	0
Software/hardware matching of management	0.022	[0.8,1]	[0.8,1]	[0.8,1]	[0,0.5]
Information release/service facilities	0.014	[0.8,1]	[0.8,1]	[0.8,1]	[0,0.5]
Operation monitor facilities	0.037	[0.8,1]	[0.8,1]	[0.8,1]	[0,0.2]
Resilience	–	0.965	0.878	–	–

$$w = (0.027, 0.096, 0.106, 0.094, 0.15, 0.117, 0.078, 0.036, 0.027, 0.04, 0.101, 0.022, 0.014, 0.037)$$

Under the condition of without emergency, the distances from resilience index of comprehensive transportation network to positive ideal solution and negative ideal solution are:

$$d_0^+ = (0, 0, 0, 0, 0.16, 0, 0.02, 0.43, 0, 0, 0, 0, 0, 0) * w^{\mathrm{T}} = 0.033$$
$$d_0^- = (1, 1, 1, 1, 0.84, 1, 0.98, 0.57, 0.8, 0.8, 1, 0.667, 0.667, 0.8) * w^{\mathrm{T}} = 0.923$$

Under the condition of with emergency, the distances from resilience index of comprehensive transportation network to positive ideal solution and negative ideal solution are:

$$\begin{aligned} d_e^+ &= (0.01, 0.04, 0.06, 0.01, 0.17, 0.5, 0.154, \\ &\quad 0.464, 0, 0, 0, 0, 0, 0) \times w^{\mathrm{T}} = 0.118 \\ d_e^- &= (0.99, 0.96, 0.94, 0.99, 0.83, 0.5, 0.846, \\ &\quad 0.536, 0.8, 0.8, 1, 0.667, 0.667, 0.8) \times w^{\mathrm{T}} = 0.849 \end{aligned}$$

Thus, the relative closeness of positive ideal solution and resilience of comprehensive transportation network without and with emergency are:

$$C(0) = \frac{0.033}{0.033 + 0.923} = 0.0346$$
$$C(e) = \frac{0.118}{0.118 + 0.849} = 0.122$$

According to the computing formula, the resilience of comprehensive transportation network without and with emergency are:

$$R(0) = 1 - C(0) = 0.965$$
$$R(e) = 1 - C(e) = 0.878$$

Namely, the resilience of comprehensive transportation network without emergency is 0.965, with emergency is 0.878. It reduced by 9.02%.

In this example, the management and service capability remains unchanged. Thus, by using above method, the network security capability declines by 12.41%.

5 Conclusion

Through considering the factors of structure capability, operation capability and management level, following the process of the status changing of the comprehensive transportation network before and after emergency till recovering to stability, from the view of security capability and management and service capability of network to structure the resilience index of comprehensive transportation network. Based on the characteristics of the qualitative and quantitative combination of the resilience metrics, propose the resilience evaluation method based on AHP and TOPSIS. Finally, using the data of Jing-Jin-Ji region comprehensive transportation network to check the method. The result proves that the proposed method is effective and practical. The result can provide technical support for improving the reliability and preventing and control the risk of the comprehensive transportation network.

Acknowledgements The authors would like to thank the Research Institute of Highway Ministry of Transport by the fundamental research project (No. 2017-9075), Qinghai science and technology project (2015-SF-A5), National Key Technologies R&D Program Strategic International Science and Technology Innovation Cooperation Projects (No. 2016YFE0206800), and the authors acknowledge the contributions of the team members.

References

1. Laprie JC (2008) From dependability to resilience. In: 38th IEEE/IFIP international conference on dependable systems and networks, Anchorage, Alaska, pp G8–G9
2. Holling CS (1973) Resilience and stability of ecological systems. Annu Rev Ecol Syst 4:1–23
3. Briguglio L, Cordina G, Farrugia N, Vella S (2006) Conceptualizing and measuring economic resilience. Constr 1:265–288
4. Huiping L, Fernandez SJ, Ganguly A (2005) Racial geography, economic growth and natural disaster resilience. House of Representative. Committee on Science and Technology, Washington, DC
5. Murray-Tuite, PM (2006) A comparison of transportation network resilience under simulated system optimum and user equilibrium conditions. In: Proceedings of the 2006 winter simulation conference WSC 06, pp 1398–1405
6. Heaslip K, Louisell WC, Collura J, Serulle NU (2010) A sketch level method for assessing transportation network resiliency to natural disasters and man-made events. In: 89th annual meeting of the transportation research board, Washington, DC
7. VTPI (2010) Evaluating transportation resilience: evaluating the transportation system's ability to accommodate diverse, variable and unexpected demands with minimal risk, TDM Encyclopedia: Evaluating Transportation Resilience. Victoria Transport Policy Institute, Victoria, BC, Canada
8. Amdal JR, Swigart SL (2010) In: Becker MC Jr (ed) Resilient transportation systems in a post-disaster environment: a case study of opportunities realized and missed in the greater New Orleans region. University of New Orleans Transportation Institute, New Orleans, LA, p 49
9. Mostashari A, Omer M, Nilchiani R (2009) Assessing resilience in a regional road based transportation network. Stevens Institute of Technology, Hoboken, NJ
10. Heaslip K, Louisell WC, Collura J (2009) Quantitative evaluation of transportation resiliency for regional networks. In: 88th annual meeting of the transportation research board, Washington, DC
11. Serulle NU, Heaslip K, Brady B, Louisell WC, Collura J (2011) Resiliency of transportation network of Santo Domingo, Dominican Republic case study. Transport Res Rec 22–30
12. Di Gangi M, Luongo A (2005) Transportation network vulnerability indicator for risk evaluation and exposure reduction. In: Proceedings of the European transport conference (ETC 2005), Strasbourg, France
13. Xia S (1995) Systematic engineering outline. Tsinghua University Press, Beijing

The Edge Importance Evaluation of Compound Network Formed by Comprehensive Transportation Network

Shuyun Niu, Jian Gao and Honghai Li

Abstract On the basis of complex network theory, the edge importance evaluation method of compound network formed by comprehensive transportation network is studied. First, combining the characteristic of comprehensive transportation network, the different network topology extraction methods are analyzed. Then the extraction method of compound network topology is determined, and the compound network of comprehensive transportation network is proposed. Second, considering different transportation ways' factors of network size, running speed and turnover volume of passenger and goods etc., the weight computation formula of every sub-network topology edge is constructed. Third, dual weighted node degree and weighted node betweenness are chosen as evaluation index which are used to evaluate the edge importance of compound network. Finally, Jing-Jin-Ji region comprehensive transportation network is chosen as empirical study. The result proves that the method proposed in this paper is feasible and effective, and the assessment results are in line with the actual situation.

Keywords Comprehensive transportation network · Compound network · Dual weighted node degree · Weighted edge betweenness · Edge importance

1 Introduction

Transportation is the important foundation of national economy and society development. It provides safe, comfortable, effective and sustainable transport service for society produce and life. Facing the ever-growing requirement of traffic travel, complicated austere nature disaster and society security situation, in case that the transportation network is blocked, the public travel, society economy running and the transportation of important goods & materials, such as electricity, coal, foodstuffs, vegetables etc. will be seriously influenced. Recent years, the serious accident

S. Niu · J. Gao (✉) · H. Li
National ITS Research Center, Research Institute of Highway Ministry of Transport, 8 Xi Tucheng Rd., Haidian District, Beijing 100088, China
e-mail: gaojian@itsc.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_16

and emergency occurs frequently in China. Such as the ice damage of 2008 in South of China, heavy snow of 2010 in North of China, which often cause regional transportation channel breaking, even wider scope traffic paralysis. The reliability of transportation network was seriously challenged.

Comprehensive transportation network normally is considered as a system, which is constituted by traffic line, port and pivot of different transportation modes cooperating with each other, supplementing mutually and coordinating closely. In the network, the traffic route works as connect line, traffic port and pivot work as junction point. All of them can transport directly or combined. Having a certain combinatorial structure and hierarchy is the specific embodiment of transportation productivity in the geographical combination. The definition shows that the comprehensive transportation has the typical network characteristic. Traffic line is edge of the network. The port and pivot are the nodes of the network. This article studies the comprehensive transportation network composed of sub-networks of highway, railway, civil aviation and waterway.

Currently, the single traffic transportation network has researched very abundantly. For example, Lewis [1] defined the key node as a node that either connected with a large number of paths, or has a significant target value. Deng [2] using the Complex Network Theory and the Communications Network Theory, integrated the characteristics of transportation networks and evaluated a highway network node importance evaluation system which included three levels of topology, shortest path, and transportation capability. Shen [3] set the functional attributes of mileage and capacity as factors when analyzing network structures and evaluated an index system for evaluating the degree of nodes and betweenness as a new method of identifying key nodes in highway networks and evaluating network capability.

Murray-Tuite [4], Heaslip et al. [5, 6], Solotani-Sobh et al. [7, 8] have studied the resilience of comprehensive transportation network. The resilience period was divided into 4 stages of normality, breakdown, self-annealing, and recovery. Wei et al., Shen et al. [9, 10], based on the complex network theory, have studied the characteristic and invulnerability of the city compound network. Li [11] have studied the compound network's vulnerability of urban agglomeration transportation network. Xu et al. [12, 13], using complex network method, studied the network topology characteristic and robustness of the compound network composed of Chinese High-speed railway and civil aviation, and the result shows that the compound network composed of high-speed railway and civil aviation and its sub-network is small world network, the robustness of compound network is better than the sub-network of high-speed railway and civil aviation.

Most of the above studies are based on the complex network theory to analysis compound network. Studies of focusing on railway and civil aviation network are more than that considering the comprehensive transportation network composed of highway, railway, civil aviation and waterway. Therefore, facing comprehensive transportation network, the method of structuring compound network is studied and the importance evaluation method of compound network edge is proposed.

2 Topology Construction of Compound Network of Comprehensive Transportation Network

2.1 Compound Network Topology Construction Method

Generally, network can be defined through three types of space method, namely space P, L and R. Space R is the network topology structure method based on route line. Spaces P and L are based on the site, but suit for different circumstance.

In the process of structuring network topology by space L method, the node is based on site. The site only has the link-edge with the adjacent sites. That's mean that the site only may influence its neighbors with a very small impact scope. This method studies only itself situation of a certain site. Such as how many other sites can be reached from the site, and which route line was taken to reach others were not involved. Therefore, scholars often call this method a docking site network topology. The disadvantage of this method is that it does not combine the sites and the route lines and the transfer situation was not mentioned.

The method of Space P analysis sites network from another point of view. In the process of structuring network topology, take the site as network node. The site not only connects with its adjacent sites, but also with all the other sites located on the same route line as itself. This method is called transfer network topology by scholars.

The method of Space R is different with Space L and P. It structures a route line network. The studied object is the relations between route lines. If there is a common site for two route lines, there will be a link-edge between them.

Take railway network as example, in the method of Space P, defining the station as node, if at least there is one train stops at two stations, the two stations connect to one edge. In the method of Space R, defining train as node, if two trains stop at least at one same station, the two trains connect to one edge. In the method of Space L, still defining station as node, if the two stations are connected by rail, and at least was taken as two adjacent stations by one running train, deem that there is an edge connecting the two stations.

This article will structure a compound network of comprehensive transportation network composed of highway, railway, civil aviation and waterway etc. Based on the transportation characteristics of highway, railway, civil aviation and waterway, and the merits and faults of above-mentioned 3 space methods, this article choose method of Space L to structure the compound network of comprehensive transportation network.

2.2 Assume and Illustration of the Compound Network Construction

Generally, the network can be abstracted into Fig. G = (V, E) composed of node group V(G) and edge group E(G). Mark the amount of node as N = |V|, the amount of edge as M = |E|. If every edge is endowed with weights, this network will be called weighted network, otherwise, called un-weighted network. Following different characteristics of different transportation modes, the undirected weighted compound network based on comprehensive transportation network will be constructed.

For convenience, 6 assumptions are given when structure the compound network of comprehensive transportation network, as follows.

(1) Choose Space L to extract the topology of compound network. Take the city where railway station or airport locates as a node. Any two cities, if there is highway connecting directly between them, or there is one same train stops at the two stations, or there is one same flight fly between them, think that there is an edge connecting the two nodes.
(2) Illustration of taking city as node. In case that a city has one or more rail stations, or civil aviation or highways passing through, the city will be abstracted into a node.
(3) Illustration of edge in compound network topology. In case that there is not only highway connecting city A and city B, but also train or flight, but not repeatedly connecting in compound network topology graph, deeming that there is only one edge connecting city A and city B. But in the process of structuring the sub-network topology graph of highway, railway and aviation, the amount of edge between nodes is equal to the amount of actual connecting highway, railway and aviation.
(4) Non-direction network. Generally, except highway, if someone can go from city A to city B by high-speed railway or flight, and also can turn back by the same way. Thus, when making the network topology extraction, neglecting the route direction, to abstract the network into non-direction network.
(5) Weighted network. Because of the different characteristics of transportation ways of highway, railway, civil aviation and waterway, there is big otherness in infrastructure scale, turnover volume of passenger-goods and transportation speed. If the edge of sub-network of different transportation way is endowed with same weighted value, the difference between different transportation cannot be reflected. Therefore, the weight factor should be considered when structuring compound network.
(6) Illustration of sub-network building. During structuring the compound network of comprehensive transportation network composed of highway, railway and civil aviation, for highway sub-network, taking prefecture-level city as node, if there is highway connecting directly between any two cities, deem that there is edge connecting between them. If there are more than one highway, deem that there are multi-edges connecting the cities. For railway sub-network, taking the city where the station locating as node, if there is one same train stops at any

two cities, deem that there is edge connecting the two cities. Similarly, for civil aviation sub-network, taking the city where the airport locates as node, if there is one same flight connects any two cities, deem that there is edge connecting the two cities. For waterway sub-network, taking the city where the port or dock locating as node, if there is one same ship line connects any two cities, deem that there is edge connecting the two cities. The sub-network is the foundation to build compound network structure.

2.3 Compound Network Constructing of Comprehensive Transportation Network

The compound network of comprehensive transportation network, composed of highway-railway-civil aviation-waterway, is constituted by node and edge. The node of the network represents prefecture-level city. If there is highway connects or train/flight stopover between two cities, deem that the two cities have edge connecting. The network composed of city nodes and edges constitutes the based architecture of compound network.

Below is the computational method of edge weight.

For highway, railway, civil aviation and waterway, because of different scale, turnover volume of passenger-goods and transportation speed, when the sub-network is modeled, the edge needs to be endowed to reflect the difference of them.

The scale, turnover volume of passenger-goods and transportation speed of the network need to be considered when the edges were endowed with weighting. Below taking highway as standard, convert other transportation ways into rate relative to highway. Namely, one edge connected by highway, the weight is 1. The weight of the edge connected by railway is the ratio value relative to highway. Through considering the above-mentioned factors, the formula to calculate the weighted value of other transportation methods is:

$$\omega_m = \frac{S_m}{S_h} \times \frac{N_m/N_h}{L_m/L_h} \times f_m \tag{1}$$

In which, ω_m is the weight of traffic mode m. For the same mode, the weight can be calculated according to the grade, such as high-speed and normal railway.

S_m, S_h are the average speeds of traffic mode m and highway.
N_m, N_h are the passenger turnover volumes of mode m and highway.
L_m, L_h are total scale of mode m and highway, respectively, namely, total mileage of infrastructure or air route.
f_m is the revise coefficient of weight of Mode m, and the range of value is 0–2.

According to the Transportation Industry Development Statistics Communiqué of 2016, the service characteristic of different transportation ways is as below. The total

mileage of railway network is 124,000 km, it mainly undertakes the transportation of long-distance passenger and goods. The annual transportation volume is 2.814 billion passengers, and the passenger turnover volume is 1257.929 billion Person kilometer. The total mileage of highway network is 4.6963 million km, it mainly undertakes the transportation of short-distance passenger and goods. The total business passenger transportation volume is 15.428 billion in 2016, and the passenger turnover volume is 1022.871 billion Person kilometer. The total mileage of inland waterway all over China is 127,100 km, it mainly undertakes the transportation of short-distance passenger, freshwater middle long distance and coastal long-distance goods. The annual transportation volume is 0.272 billion passengers, and the passenger turnover volume is 7.233 billion Person kilometer. The total mileage of civil aviation route is 6.348 million km, it mainly undertakes the transportation of long-distance passenger. The annual transportation volume is 0.488 billion passengers, and the passenger turnover volume is 835.954 billion Person kilometer.

According to the above-mentioned formula of weight, and the statistical data of 2016 [14] in case of $f_m = 1$, the weight of topology graph edge of diversified transportation ways relative to the edge of highway is:

$$\omega_{GR} = 7.0,\ \omega_{HR} = 14.0,\ \omega_W = 0.01,\ \omega_A = 0.48$$

In which, ω_{GR}, ω_{HR}, ω_W, ω_A, represent the weights of general railway, high-speed railway, waterway and civil aviation, respectively.

3 The Edge Importance Evaluation of Compound Network

3.1 The Evaluation Index Selection

In the process of importance evaluation to compound network edge, choose dual weighted node degree and weighted edge betweenness to make comprehensive evaluation. The dual weighted node degree can reflect the local importance of the edge. The betweenness can reflect the importance of the edge for connection in the whole network. That is, the weighted edge betweenness can reflect the global importance of the edge.

Through dual topology method, abstract the topology graph into node, called dual nodes. The connecting relation between edges, namely node, is abstracted into edge. In case there is node connects the two edges, deem there is edge connecting between the two dual nodes.

Dual weighted node degree was gotten through dual topology graph of compound network topology. Weighted edge betweenness was gotten through compound network topology graph calculation based on Space L extraction.

(1) **Dual weighted node degree**

Dual weighted node degree of dual node i is equal to the sum of weight of the dual node connecting directly with dual node i. It reflects the local importance of the dual node i, namely, local importance of the edge in original topology. The bigger the value is, the more important the local position of the edge in the network is. The calculation formula of the dual weighted node degree is:

$$D_i = \sum_m D_i^m = \sum_m \sum_j \omega_{ij}^m \tag{2}$$

D_i is dual weighted node degree for dual node i.
D_i is the dual weighted node degree of transportation mode m for node i.
ω_{ij}^m is the weight of dual node j connecting with dual node i for transportation mode m.

(2) **Weighted edge betweenness**

In the compound network, the betweenness is the ratio of amount of shortest route passing through node i and amount of shortest route among all the nodes. The bigger the weighted node betweenness is, the more important the connectivity of node to the whole of network is. The formula is:

$$b_e = \frac{\sum_{j,k} n_{jk}(e)}{\sum_{j,k} n_{jk}} \tag{3}$$

In which, i, j, k are nodes identification.

b_e is weighted node betweenness of node i.
$n_{jk}(e)$ is the number of shortest routes connecting node j and node k passing through edge e.
n_{jk} is the number of shortest routes connecting node j and node k.

3.2 The Edge Importance Evaluation Model

The importance evaluation of compound network edge of comprehensive transportation network mainly constitutes 7 steps. (1) Structure sub-network topology model of diversified transportation ways based on Space L. Calculate the edge weight of every network topology graph separately. (2) Structure compound network topology model of comprehensive transportation network based on Space L. Calculate the edge weight and mileage of compound network. (3) Extract the dual topology graph of compound network. (4) Calculate the dual weighted node degree and weighted edge betweenness. (5) Standardize the dual weighted node degree and weighted edge

betweenness. (6) Determine the weight of evaluation index. (7) Calculate the edge importance of compound network.

The concrete procedure is as below:

Step 1: **Data preparation**

Enter the data of highway, railway and civil aviation network separately. Use the method of Space L to extract every sub-network topology graph by taking city as node. Fix the edge mileage and weight of every sub-network topology.

Step 2: **Create sub-network adjacent matrix based on sub-network topology graph**

Create the weighted adjacent matrix and mileage adjacent matrix of every sub-network based on the connecting relation and the mileage & weight of sub-network edge following step 1.

Step 3: **Extract the topology structure of compound network**

On the base of every sub-network topology, build the topology structure model of compound network. The mileage of edge is fixed through the middle position of mileage of sub-network edge. The weight of edge is the weight sum of every sub-network edge between two nodes. Accordingly, create the mileage adjacent matrix and weight adjacent matrix of compound network.

Step 4: **Dual topology graph extraction of compound network**

Extract the dual topology graph of compound network based on the original topology graph of compound network. Abstract the edge of original topology graph into dual node.

Step 5: **Calculate the dual weighted node degree and weighted edge betweenness of compound network**

Through weighted adjacent matrix, calculate the weighted node degree of every dual node. Through mileage adjacent matrix, using Dijkstra algorithm, calculate the weighted mileage betweenness of every edge.

Step 6: **Standardize the parameters**

Standardize the index of dual weighted node degree and weighed edge betweenness. The formula is:

$$x_s = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{4}$$

In which, x_s is the standard value of the parameter.
$x_{\max}$, $x_{\min}$ are the maximum and minimum value of the parameter sequence.

Step 7: **Determine the weight of evaluation index**

Through expert marking method, fix the weight value of dual weighted node degree and weighed edge betweenness.

Step 8: **Calculate the edge importance of compound network**

Based on above results, adopting weighted mean method, calculate the importance of compound network edge.

4 Analysis on Jing-Jin-Ji Region Comprehensive Transportation Network

The Beijing-Tianjin-Hebei region is called Jing-Jin-Ji region for short. In the Jing-Jin-Ji region, both the road network and the railway network are relatively developed. They consist of radiation and horizontal and vertical lines centered on Beijing. The highway network has been formed, trunk lines and branch roads are connected to each other and lead in all directions. The railway network consists of a number of radiations from Beijing and a number of north-south vertical and horizontal lines. The railway itself has also formed a network. The area of Jing-Jin-Ji region is 218,000 km^2, and there are 6 aeronautical stations in Jing-Jin-Ji Region. But because of the short space distance and developed highway and railway networks, there are only 3 direct flights (Shijiazhuang-Qinhuangdao,Shijiazhuang-Tianjin, Shijiazhuang-Zhangjiakou). There are 4 ports in the region composed of Huanghua, Tianjin, Jingtang and Qinhuangdao. All of them locate at southeast coastal area and connecting with inland by highway and railway. And their main business is goods transportation. Therefore, the main comprehensive transportation networks are highway, railway and civil aviation in this area. The railway and civil aviation take city as the hub node, their stations are connected by highway, and the daily operation of them is independent. Therefore, this section firstly analysis the highway and railway network. The edge importance of compound network of highway, railway and civil aviation will be evaluated afterward.

(1) **Analysis of the highway and railway network of Jing-Jin-Ji region**

According to the database of highway foundation, extract the connecting relation matrix and mileage adjacent matrix of highway network in Jing-Jin-Ji region. Using UCINET and Matlab to analysis, the connecting relation topology graph of highway network can be gotten as Fig. 1.

Based on the railway network data of 2016 in the area, extract railway network topology graph is as Fig. 2.

According to above evaluation method of edge importance, the importance value of every section can be calculated. Table 1 is the Top 5 sections of importance ranking of highway and railway networks in the area.

According to Table 1, in the highway and railway network of Jing-Jin-Ji region, there is big differences among the edges of high importance. In every network, the top 5 edges are different. It reflects the principle of mutual supplement and support among different transportation ways, and proves that it's meaningful to study the importance evaluation of compound network of comprehensive transportation network.

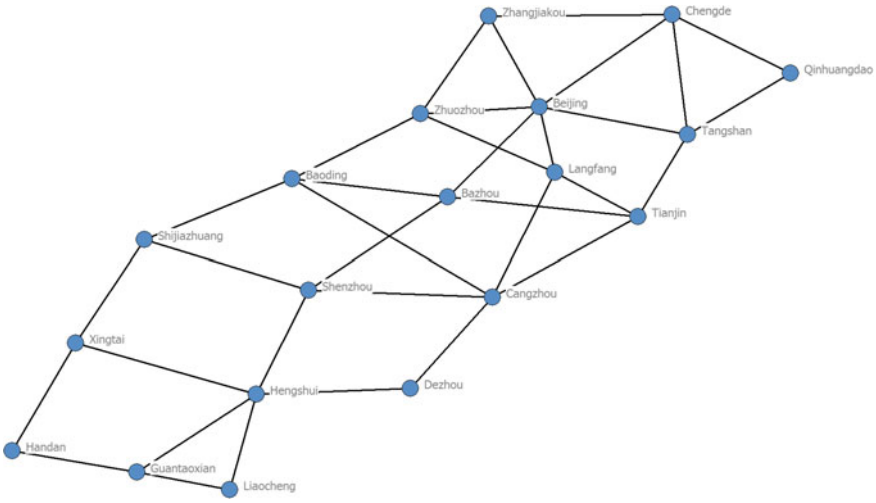


Fig. 1 Topology graph of highway network of Jing-Jin-Ji region

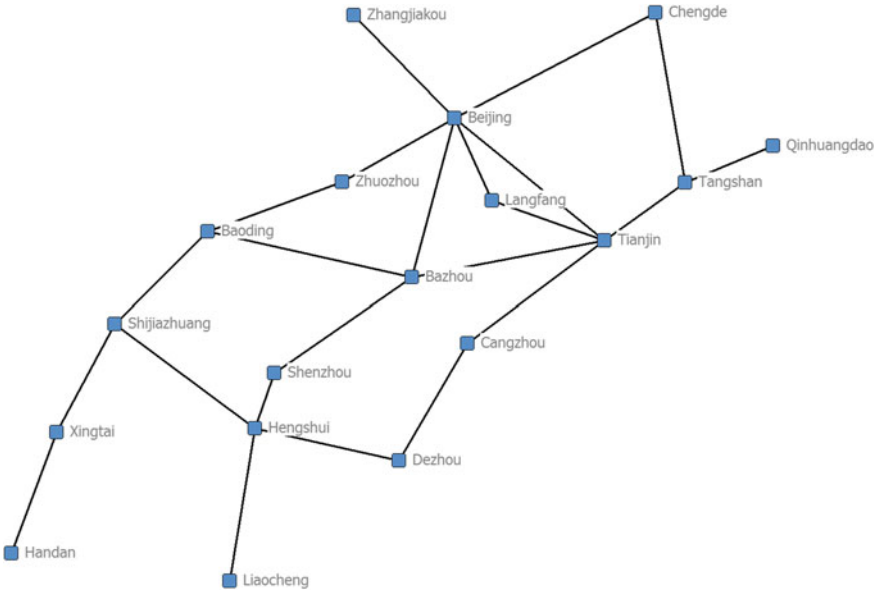


Fig. 2 Topology graph of railway network of Jing-Jin-Ji region

Table 1 Top 5 section of importance ranking

	Ranking	Origin	Destination	Standardized weighted edge betweenness	Standardized dual weighted node degree	Evaluation result of edge importance
Highway network	1	Shenzhou	Hengshui	1.00	0.33	0.80
	2	Tianjin	Cangzhou	0.65	0.75	0.68
	3	Tangshan	Tianjin	0.63	0.75	0.67
	4	Cangzhou	Shenzhou	0.70	0.42	0.61
	5	Zhuozhou	Baoding	0.57	0.67	0.60
Railway network	1	Tangshan	Tianjin	1.00	0.63	0.89
	2	Baoding	Shijiazhuang	1.00	0.44	0.83
	3	Tianjin	Cangzhou	0.92	0.59	0.82
	4	Beijing	Tianjin	0.81	0.39	0.68
	5	Tianjin	Bazhou	0.72	0.54	0.67

(2) **Edge importance evaluation of compound network of highway-railway-civil aviation**

According to the topology structure of above highway and railway network, through superposition, the connecting relation graph of compound network of highway-railway-civil aviation can be shown in Fig 3.

In Fig. 3, the dotted lines among Tianjin, Shijiazhuang, Qinhuangdao and Zhangjiakou are topology edge of civil aviation network.

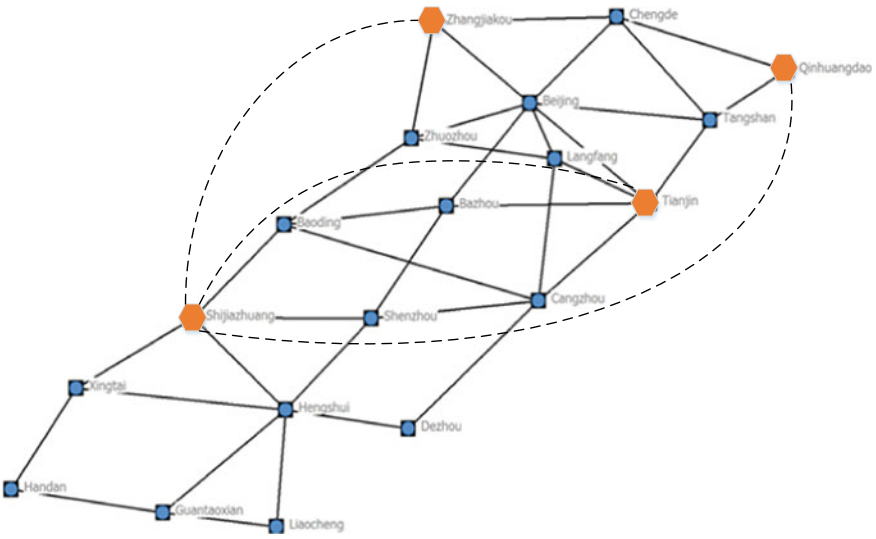


Fig. 3 Connecting relation graph of compound network of highway-railway-civil aviation

Based on the method of edge importance evaluation described in Sect. 3.2, combining method of expert marking, fix the weights of dual weighted node degree and weighted edge betweenness are 0.3 and 0.7. Namely, during the comprehensive evaluation to edge importance, think the weight of global importance is higher than the weight of local importance. The calculation results of edge importance of compound network are as Table 2.

The discrete coefficients of the dual weighted node degree and weighted edge betweenness are 0.47 and 0.93 in the compound network of Jing-Jin-Ji region, respectively. It shows that the discrete coefficient of weighted edge betweenness is higher, and means that there is a big difference between the edge global importance of compound network in the comprehensive transportation network of Jing-Jin-Ji region. Some few edges undertake the bridge function of most network nodes connecting.

Distribution curve of dual weighted node degree and weighted edge betweenness result is as Fig. 4.

In Fig. 4, lateral axis is the number of edge in compound network topology. Vertical axis is the standard values of dual weighted node degree, weighted edge betweenness and edge importance.

It is known from Table 2 to Fig. 4.

(1) Top 5 ranking of weighted edge betweenness are Shenzhou-Hengshui, Tianjin-Cangzhou, Cangzhou-Shenzhou, Baoding-Shijiazhuang, and Tangshan-Tianjin. That means these 5 edges play pivotal role in the comprehensive transportation network of Jing-Jin-Ji region. It has global importance and was passed through many times by the shortest route. The first ranking of weighted edge betweenness is Shenzhou-Hengshui. It is connected by highways and railways and plays a pivotal role to connect South and North network.
(2) Top 5 ranking of dual weighted node degree are Beijing-Tangshan, Beijing-Zhuozhou, Tangshan-Tianjin, Tianjin-Shijiazhuang and Tianjin-Cangzhou, That means these 5 edges have greater local connectivity and local importance. All the connected nodes are large pivotal cities.
(3) Top 5 ranking of evaluation result of importance are Shenzhou-Hengshui, Tianjin-Cangzhou, Tangshan-Tianjin, Baoding-Shijiazhuang and Cangzhou-Shenzhou. It reflects these 5 edges have both global and local importance and coincide with actual situation.

5 Conclusion

On the basis of complex network theory, this article studies the importance evaluation method of compound network structure and edge of comprehensive transportation network. Combining the characteristic of comprehensive transportation network, propose the extraction method of compound network topology structure. By considering different transportation ways' network size, running speed and turnover volume of passenger-goods etc. propose the weight computation formula of every sub-network

Table 2 Evaluation result of edge importance

No.	Origin	Destination	Standard value of dual weighted node degree	Standard value of weighted edge betweenness	Evaluation result of edge importance
1	Shenzhou	Hengshui	0.23	1.00	0.77
2	Tianjin	Cangzhou	0.73	0.71	0.72
3	Tangshan	Tianjin	0.77	0.69	0.71
4	Baoding	Shijiazhuang	0.59	0.69	0.66
5	Cangzhou	Shenzhou	0.32	0.76	0.63
6	Zhuozhou	Baoding	0.55	0.62	0.60
7	Shijiazhuang	Xingtai	0.53	0.52	0.52
8	Xingtai	Handan	0.38	0.55	0.50
9	Beijing	Bazhou	0.73	0.38	0.48
10	Beijing	Zhuozhou	0.78	0.31	0.45
11	Hengshui	Guantaoxian	0.14	0.55	0.43
12	Qinhuangdao	Tangshan	0.40	0.40	0.40
13	Bazhou	Shenzhou	0.28	0.40	0.37
14	Hengshui	Xingtai	0.40	0.31	0.34
15	Beijing	Tianjin	1.00	0.05	0.33
16	Beijing	Tangshan	0.82	0.12	0.33
17	Beijing	Langfang	0.69	0.17	0.32
18	Tianjin	Bazhou	0.68	0.14	0.30
19	Guantaoxian	Liaocheng	0.00	0.43	0.30
20	Langfang	Cangzhou	0.42	0.24	0.29
21	Cangzhou	Dezhou	0.40	0.24	0.29
22	Bazhou	Baoding	0.50	0.17	0.27
23	Tianjin	Langfang	0.65	0.10	0.26
24	Baoding	Cangzhou	0.55	0.10	0.23
25	Tianjin	Shijiazhuang	0.77	0.00	0.23
26	Langfang	Zhuozhou	0.43	0.14	0.23
27	Shijiazhuang	Shenzhou	0.36	0.17	0.23
28	Beijing	Chengde	0.64	0.02	0.21
29	Hengshui	Dezhou	0.31	0.17	0.21
30	Handan	Guantaoxian	0.13	0.24	0.20
31	Shijiazhuang	Hengshui	0.45	0.10	0.20
32	Beijing	Zhangjiakou	0.60	0.02	0.20
33	Chengde	Tangshan	0.41	0.10	0.19
34	Zhangjiakou	Zhuozhou	0.34	0.12	0.19

(continued)

Table 2 (continued)

No.	Origin	Destination	Standard value of dual weighted node degree	Standard value of weighted edge betweenness	Evaluation result of edge importance
35	Qinhuangdao	Shijiazhuang	0.40	0.00	0.12
36	Zhangjiakou	Shijiazhuang	0.37	0.00	0.11
37	Chengde	Qinhuangdao	0.23	0.02	0.08
38	Zhangjiakou	Chengde	0.20	0.02	0.08
39	Hengshui	Liaocheng	0.18	0.00	0.05

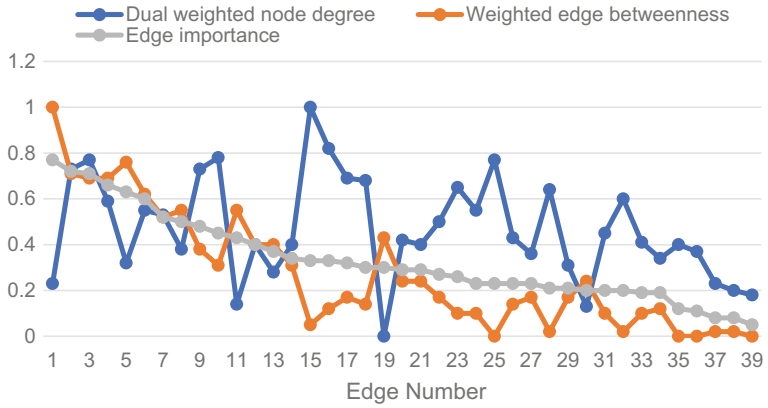


Fig. 4 Distribution of evaluation result of edge importance in compound network

topology edge. Choose dual weighted node degree and weighted node betweenness as evaluation index to structure the important evaluation method facing compound network edge of comprehensive transportation network. Finally, take Jing-Jin-Ji region comprehensive transportation network as example to verify the proposed method. The result matches the reality well. And the importance of the compound network can be reflected completely. The result of this article can effectively support the reliability improvement and safety risk prevention of comprehensive transportation network.

Acknowledgements The authors would like to thank the Research Institute of Highway Ministry of Transport by the fundamental research project of "Research on the Development Roadmap of China Intelligent and Network Transportation System" (No. 2017-9075), and the authors acknowledge the contributions of the team members.

References

1. Lewis TG (2011) Network science theory and applications. China Machine Press
2. Deng Y, Yang Y, Ma R (2010) Highway network structure characteristics based on complex network theory. China J Highw Transp 23:98–104
3. Shen H (2012) An approach to analysis and evaluation of highway network structural properties for risk assessment and emergency management. Beijing Jiao Tong University Doctoral Dissertion
4. Murray-Tuite PA (2006) Comparison of transportation network resilience under simulated system optimum and user equilibrium conditions. Presented at winter simulation conference
5. Heaslip K, Louisell WC, Collura J (2009) Quantitative evaluation of transportation resiliency for regional networks. In: 88th annual meeting of the transportation research board, Washington, DC
6. Heaslip K, Louisell WC, Collura J, Serulle NU (2010) A sketch level method for assessing transportation network resiliency to natural disasters and man-made events. In: 89th annual meeting of the transportation research board, Washington, DC
7. Soltani-Sobh A, Heaslip K, El Khoury J (2015) Estimation of road network resiliability on resiliency: an uncertain based model. Int J Disaster Risk Reduct 14:536–544
8. Soltani-Sobh A, Heaslip K, Stevanovic A, El Khoury J, Song Z (2016) Evaluation of transportation network reliability during unexpected events with multiple uncertainties. Int J Disaster Risk Reduct 17:128–136
9. Wei Z, Gan Y, Zhao P (2015) Several characteristic studies of city complex transportation network. Eng Inf Transport Syst 15(1):106–111
10. Shen L, Zhang D, Xiang Y, Wang Z, Zhang T (2017) The simulation study of invulnerability and cascading failure to complex network of city metro and bus. J Southwest Jiao Tong Univ 52(6):1–9
11. Li C, Wei L, Li FX, Zheng L (2017) The study of frangibility of urban agglomeration complex network based on attack strategy. Road Traffic Sci Technol 34(3):101–109
12. Xu F, Zhu J, Yang W (2013) Structure of complex network and analysis to network topology characteristic of high speed railway and civil aviation. Complicated Syst Complex Sci 10(3):1–11
13. Xu F, Zhu J, Miao J (2015) Robustness study of aeronautical railway complex network based on complicated network. Complicated Syst Complex Sci 12(1):40–45
14. The Transportation Industry Development Statistics Communiquéé of 2016

Economic Benefit Optimization Model of Urban Rail Station Based on C-D Function

Hua-lan Wang, Jia-ying Xu, Zun-jie Hu and Man Li

Abstract In order to study the optimal investment plan for urban rail stations under a certain amount of investment, the safe investment is regarded as a production process, "safe value-reduced output" and "safe value-added output" after the reduction of accidents are as benefits and the Cobb-Douglas function have been improved. Employee safe training investment, safe equipment maintenance investment, safe publicity and guidance logo investment, emergency evacuation drilling investment, and security staff salary input have been taken as independent variables to build safe investment in the stations to optimize the model. Taking an example in Shenzhen to verify the type by use of Matlab. The results show that the model is a good reflection of the relationship between the safe investment sub-items of urban rail stations and the economic losses of accidents during the ten years, and the benefits of the optimized investment plan are obvious.

Keywords Cobb-Douglas production function · Accident economic loss · Safe investment economic benefit · Measurement model

1 Introduction

With the acceleration of urbanization, the construction of urban rail transit in China has developed rapidly. By the end of 2017, a total of 165 urban rail lines have been operated in 34 cities. Urban rail transit provides people with quick and comfortable travel services by its unique operating characteristics but has more prominent safe problems than ordinary ground transportation because of its semi-closed underground space. Because of the dense passenger flow and its difficult decentralization, the urban rail station is a major potential dangerous point. The safety of stations is significant to the safe operation of the entire urban rail transit system. It is an important way to improve the station safety that increasing safe investment in station equipment, management, operation, maintenance and repair. However, safe investment

H. Wang (✉) · J. Xu · Z. Hu · M. Li
School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China
e-mail: wanghualan126@126.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_17

has a limitation whether considering the benefits of the company or the effect of safe investment. With certain investment money, different investment plans will produce different benefits. Considering the characteristics of urban rail station investment, we use the improved Cobb-Douglas production function to study the quantitative relationship between safe investment and safe benefits of rail transit stations. We hope the study can provide the help for decision-making on station security invest.

Research on the safety of urban rail transit stations focuses on security evaluation [1] and passenger evacuation in emergency [2]. There are few researches on safe investment.

The Cobb-Douglas function (referred to as C-D function) is widely used form of production function in economics. It expresses the quantitative relationship between output and input labor and capital [3]. It has been applied in many fields [4–6]. The paper regards the safe investment of the station as a production process, the"safe value-reduced output" and "safe value-added output"is regarded as revenue generated and the C-D function is improved, employee safe training investment, safe equipment maintenance investment, safe propaganda and guidance logo investment, emergency evacuation drill investment, and salaries of security personnel were independent variables, and the safe investment production function of urban rail stations was constructed to quantitatively express the relationship between safe investment and safe benefits.

The urban rail station is a new type of industry. The economic benefits of its safe investment are different from other production activities. Although the C-D production function has been widely used and the theoretical basis of the method has been shaped, it is aimed at the special project of the urban rail station are few studies, it is necessary to establish a metropolitan-station safe investment economic benefit measurement model that considers the factors of safe investment and that the parameters have good maneuverability in theory and practical application.

2 Model Establishment

Urban rail station safety refers to a state in which the possible damages to human life, property, and environment caused by the operation of the system can be reduced to an acceptable level while maintaining normal transportation production. The primary way is to invest in safety facilities, emergency rescue drills, and employee safety training. The safe investment of urban rail stations is defined as a series of investment and operational activities to ensure the efficient and safe implementation of urban rail transit passenger operations, and put part of the costs into the safe area of the stations. The safe investment mainly includes employee safe training, safe equipment maintenance, safe promotion and guidance signs, emergency evacuation drills, and security staff wages.

Safe investment benefits are divided into economic benefits and non-economic benefits. Safe economic benefits include safe value-added benefits and safe value-reduced benefits. Safe value-added benefits refer to the maintenance of normal production and operations of enterprises through safe investments, reducing the impact of accidents on the operation of enterprises, so as to achieve the value-added benefits. The safe value-reduced benefits refer to the safety inputs. The protection will reduce the accident rate, which will directly reduce casualties and economic losses caused by the accident. The non-economic benefits of safety mainly refer to the positive effects that security investment has in social security, stability and environmental harmony. This paper only studies the safe and economic benefits.

To facilitate the establishment of the model assume that:

(1) The total investment during the operation period of the urban rail station is divided into productive investment and safe investment.
(2) The safe and economic benefits are equal to the safe outputs, expressed as the sum of value-reduced output and value-added output. The difference value in the economic losses of safe accidents before and after the optimization of safe investment for value-reduced output is expressed, and the difference value of the change in the output value of the enterprise before and after optimization of safe investment for value- added is expressed.
(3) The total amount of safe investment before and after the optimization of station safe investment is the same. The difference lies in the proportion of inputs for safe sub-items.

2.1 Accident Economic Loss Model

In the safe investment of urban rail stations, the investment of each safe factor is independent, and the accidental economic loss curve and the input-output curve of the C-D production function are all the non-linear curve [7]. The relationship between it and the accident loss is related to the input and output of C-D production function which is very consistent. C-D production function can more closely simulate the functional relationship between the two. According to the C-D production function [8], the safety output of urban rail stations is a function of the safety inputs, which is:

$$L = F(p_1, p_2, p_3, p_4, p_5) = A p_1^{a_1} p_2^{a_2} p_3^{a_3} p_4^{a_4} p_5^{a_5} \tag{1}$$

In the formula:

L is the economic loss of safe accidents;

P_1, P_2, P_3, P_4, P_5 $P_i > 0$ ($i = 1, 2, \ldots, 0.5$) are investment in safe equipment maintenance, investment in safe propaganda and guidance, investment in emergency evacuation drills, investment in security personnel, and investment in employee safe training.

A is a constant and represents a technical element.

a_1, a_2, …, a_5 are the output elasticity factors of safe investment factors in safe production respectively.

As the accident loss cannot be negative, that is $L = Ap_1^{a_1} p_2^{a_2} p_3^{a_3} p_4^{a_4} p_5^{a_5} > 0$, $A > 0$, it can be seen that if the safe sub-item investment amount increases, then the safe accident loss decreases, then:

$$\frac{\partial F}{\partial p_i} = a_i \frac{L}{p_i} < 0 \Rightarrow a_i < 0; \lim_{p_i \to 0} L = \lim_{p_i \to 0} A p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_n^{\alpha_n} = \infty$$

In this formula, the investment amount for each safe sub-item cannot be non-zero, which proves that the accident economic loss model represented by the C-D production function satisfies the conditions that should be satisfied.

Take the logarithm of both sides of the Formula (1):

$$\ln L = \ln A + a_1 \ln p_1 + a_2 \ln p_2 + a_3 \ln p_3 + a_4 \ln p_4 + a_5 \ln p_5 \tag{2}$$

The total amount of economic expenditure required for safety is called the safe burden function, including total safe investment and economic losses caused by accidents, which is:

$$P = \sum_{i=1}^{n} p_i \tag{3}$$

The security burden function is:

$$B = L + P = L + \sum_{i=1}^{n} p_i \tag{4}$$

The goal of safe investment in stations is to ensure the safety of the station under certain conditions with the minimum security burden. Therefore, the paper establishes the following investment optimization model with the objective of minimizing the safe burden like:

$$\min B = L + \sum_{i=1}^{n} p_i$$

restrictions:

$$\mathrm{L} = \mathrm{F}(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_5) \leq l$$

$$P = \sum_{i=1}^{n} p_i \geq m$$

$$0 \leq p_i \leq k_i \ (i = 1, 2, \ldots, 5)$$

In the formula, the first constraint is that the accident loss of the urban rail station does not exceed the threshold value l; the second constraint is that the total amount of safe investment cannot be lower than the minimum value m; the third constraint is the non-zero and maximum value of the investment of each safe sub-item constraint.

In solving such a nonlinear model with constraints, the optimization function Fmincon of the Optimization Toolbox in MATLAB can be used for solving.

2.2 Security Value-Reduced Output Model

Safe value-reduced refers to the difference between accident losses before and after optimization of safe investment optimization.

Before the optimization of safe investment, under the premise of a certain total amount of safe investment, historical data are usually used to analyze its investment structure, calculate the investment proportion of each safe sub-item and solve the safe sub-items in the coming years. Using the regression analysis method to obtain the values of the parameters in the model, the results are substituted into the model of accidental economic loss at the station (Formula 3), and the expected direct economic losses are calculated as:

$$L_0 = A^* p_{10}^{\alpha_1} p_{20}^{\alpha_2} \cdots p_{n0}^{\alpha_n} \tag{5}$$

The direct economic loss after using the safe investment optimization model is:

$$L^* = A^* p_1^{*\alpha_1} p_2^{*\alpha_2} \cdots p_n^{*\alpha_n} \tag{6}$$

The security value-reduced output can be calculated based on the difference between before and after safe investment optimization, expressed as:

$$F_1 = L_0 - L^* = A^* p_{10}^{\alpha_1} p_{20}^{\alpha_2} \cdots p_{n0}^{\alpha_n} - A p_1^{*\alpha_1} p_2^{*\alpha_2} \cdots p_n^{*\alpha_n} \tag{7}$$

L* is accidental economic loss after optimization of safe investment optimization.
L_0 is accidental economic loss before optimization of safe investment optimization.

2.3 Security Value-Added Output Model

There is a correlation between the annual output value of the station and the situation of productive investment and safe accidents. Reference [9] establishes an annual production value model for the station as:

$$V = bX^{\eta_0} \mathrm{e}^{\eta L} \tag{8}$$

In the formula:

V is annual output value of station.
X is productive investment of station.
L is economic loss of accident.
b, η_0, η are all coefficient.

Take the logarithm on both sides of the Formula (8):

$$\ln V = \ln b + \eta_0 \ln X + \eta L \tag{9}$$

Using the regress function in MATLAB, combined with the station output value, accident losses and investment values over the years, using multiple linear regression to complete the equation.

Expression F_2 of safe value-added output:

$$F_2 = V^* - V_0 = bX^{\eta_0} e^{\eta L^*} - bXe^{\eta L_0} \tag{10}$$

V* is annual or estimated annual production value of urban rail stations after optimizing safe investment.
V_0 is the estimated annual production value under the original investment.

2.4 Safe Investment Economic Benefit Model

According to the model's assumption 2, the expression of safe output is:

Safe Output F = Safe Value-Reduced Output F_1 + Safe Value-Added Output F_2.

$$\begin{aligned} F &= F_1 + F_2 \\ &= \left(A^* p_{10}^{\alpha_1} p_{20}^{\alpha_2} \cdots p_{n0}^{\alpha_n} - A p_1^{*\alpha_1} p_2^{*\alpha_2} \cdots p_n^{*\alpha_n}\right) \\ &\quad + \left(bX^{\eta_0} e^{\eta L^*} - bXe^{\eta L_0}\right) \end{aligned} \tag{11}$$

Station safe benefits can be expressed as the difference between output and inputs as:

$$\begin{aligned} E &= F - P \\ &= F_1 + F_2 - P \\ &= \left(A^* p_{10}^{\alpha_1} p_{20}^{\alpha_2} \cdots p_{n0}^{\alpha_n} - A p_1^{*\alpha_1} p_2^{*\alpha_2} \cdots p_n^{*\alpha_n}\right) + \left(bX^{\eta_0} e^{\eta L^*} - bXe^{\eta L_0}\right) - P^* \end{aligned} \tag{12}$$

Can also be used to express the ratio of safe output to safe input as:

$$\begin{aligned} E &= F/P \\ &= (F_1 + F_2)/P \\ &= \left(A^* p_{10}^{\alpha_1} p_{20}^{\alpha_2} \cdots p_{n0}^{\alpha_n} - A p_1^{*\alpha_1} p_2^{*\alpha_2} \cdots p_n^{*\alpha_n}\right) + \left(bX^{\eta_0} \mathrm{e}^{\eta L^*} - bX\mathrm{e}^{\eta L_0}\right)/P^* \end{aligned} \quad (13)$$

3 Case Analysis

Table 1 shows the safe investment sub-items, accident losses, production investment amount and annual output value of a station in Shenzhen in the past 10 years. It is a visual display of the relationship between safe investment benefits. Assume that the station's annual revenue is owned by the station, which is the total annual output value of the station.

Calculate the logarithm of each year's safe investment sub-items and total accident losses. Use the regress function in the MATLAB software to perform the regression calculation, which are:

$$\begin{aligned} &\ln A^* = 4.4838;\ \alpha_1 = -0.0633;\ \alpha_2 = -0.0855;\ \alpha_3 = -0.4561; \\ &\alpha_4 = -0.0437;\ \alpha_5 = -0.1621 \\ &\ln L = 4.4838 - 0.0633 \ln p_1 - 0.0855 \ln p_2 \\ &\qquad - 0.4561 \ln p_3 - 0.0437 \ln p_4 - 0.1621 \ln p_5 \\ &L = 88.5706 p_1^{-0.0633} p_2^{-0.0855} p_3^{-0.4561} p_4^{-0.0437} p_5^{-0.1621} \end{aligned}$$

When using the statistical toolbox in MATLAB to solve the problem, the significance probability is found to be less than 0.05, so the null hypothesis is rejected, and it is concluded that at least one independent variable in the equation is not equal to zero. It is known from the literature [10] that the regression equation is significant. In addition, due to $R^2 = 0.9415$, the sum of the squares of the coefficient errors is small, the proof model can well reflect the relationship in the past ten years, and the fitting degree is high.

The model was used to optimize the safe investment of the station. The accident loss threshold l was taken as 250,000 yuan; the minimum total safe investment m was taken as 100,000 yuan. Using the fmincon function in the MATLAB software to solve the problem, the investment amount for each safe item under the optimized investment plan is:

$$\begin{aligned} &p_1^* = 1.4408,\ p_2^* = 1.9461, \\ &p_3^* = 10.3816,\ p_4^* = 0.9947,\ p_5^* = 3.6898 \end{aligned}$$

Table 1 Investment and output values over the years of the station (10,000 yuan)

Years	Safe facilities maintenance	Safe promotion guide logo	Emergency evacuation drill	Safe payroll	Employee safe training	Safe investment	Total accident loss	Production investment	Gross output value
1	27.88	1.32	3.36	18.44	36.59	87.95	18.50	1003.66	1603.59
2	31.49	0.71	2.98	21.03	8.43	64.64	29.00	1028.58	1685.81
3	20.64	0.25	1.64	13.41	10.98	46.92	49.55	1432.40	1958.63
4	36.19	0.55	2.17	29.76	1.93	70.60	34.58	1505.63	2186.37
5	28.54	0.40	1.98	17.04	2.25	50.21	46.31	1742.89	2617.57
6	23.46	0.95	1.59	13.65	2.30	41.95	42.29	1974.29	2835.57
7	31.63	2.80	1.56	21.93	2.42	60.34	44.63	1997.62	3229.23
8	49.59	2.74	0.65	27.86	2.48	83.32	59.64	2033.56	3361.02
9	31.23	3.12	0.68	17.16	4.04	56.23	49.12	2175.44	3533.34
10	21.74	1.18	0.73	31.19	2.69	57.53	61.72	2239.68	3604.72

The accidental economic loss L^* after the optimization:

$$\begin{aligned} L^* &= 88.5706 p1^{-0.0633} p2^{-0.0855} p3^{-0.4561} p4^{-0.0437} p5^{-0.1621} \\ &= 88.5706 \times 1.4408^{-0.0633} 1.9461^{-0.0855} 10.3816^{-0.4561} 0.9947^{-0.0437} 3.6898^{-0.1621} \\ &= 22.7613 \text{ (10000 yuan)} \end{aligned}$$

The total amount of safe investment remains unchanged. Based on the percentage of each sub-item in the investment data over the years, the investment amount for each safe sub-item before the projected annual safe investment optimization is obtained p_{i0}:

$p_{10} = p_0 * 49.41\% = 18.453 * 49.41\% = 9.12$ (Ten thousand yuan);
$p_{20} = p_0 * 2.23\% = 18.453 * 2.23\% = 0.41$ (Ten thousand yuan);
$p_{30} = p_0 * 2.85\% = 18.453 * 2.85\% = 0.53$ (Ten thousand yuan);
$p_{40} = p_0 * 34.54\% = 18.453 * 34.54\% = 6.37$ (Ten thousand yuan);
$p_{50} = p_0 * 10.97\% = 18.453 * 10.97\% = 2.02$ (Ten thousand yuan).

Accident economic losses before optimization of safe investment:

$$\begin{aligned} L_0 &= 88.5706 p_{10}^{-0.0633} p_{20}^{-0.0855} p_{30}^{-0.4561} p_{40}^{-0.0437} p_{50}^{-0.1621} \\ &= 88.5706 \times 9.12^{-0.0633} 0.41^{-0.0855} 0.53^{-0.4561} 6.37^{-0.0437} 2.02^{-0.1621} \\ &= 91.3584 \end{aligned}$$

The security value-reduced output is:

$F_1 = L_0 - L^* = 91.3584 - 22.7613 = 68.5971$ (Ten thousand yuan).

Using MATLAB software to perform linear regression calculations on the logarithm of production investment, the total accident loss, and the logarithm value of the total output value:

$$\begin{aligned} &\ln b = 0.0484; \eta_0 = 1.0556; \eta = -0.0008 \\ &\ln V = 0.0484 + 1.0556 \ln X - 0.0008 L \\ &V = 1.0496 X^{1.0556} \mathrm{e}^{-0.0008x} \end{aligned}$$

Significant probability $P = 0.0000$, rejecting the null hypothesis, considers the function to be meaningful, and can better reflect the functional relationship between the economic losses of security accidents and the safe sub-item investment of the station in the past ten years.

Assume that the station's investment for operations is 12 million yuan, and the expected production value V0 is:

$$\begin{aligned} V_0 &= 1.0496 X^{1.0556} \mathrm{e}^{-0.0008 * 91.3584} \\ &= 1.0496 \times 1200^{1.0556} \mathrm{e}^{-0.0008 * 91.3584} \\ &= 1736.4664 \end{aligned}$$

The output value V^* after optimizing investment is:

$$\begin{aligned} V* &= 1.0496X^{1.0556}\mathrm{e}^{-0.0008*L^*} \\ &= 1.0496 \times 1200^{1.0556}\mathrm{e}^{-0.0008*22.7613} \\ &= 1834.4229 \end{aligned}$$

$$F_2 = V^* - V_0 = 97.9565$$

Safe and economic benefits:

$$\begin{aligned} E &= F - P \\ &= (F_1 + F_2) - P = 148.1006 \end{aligned}$$

Analysis of results: The safe and economic benefit amounted to approximately 1.48 million yuan, which was far greater than the safe investment amount P (about 180,000 yuan), confirming the feasibility of the economic model for the economic investment in urban rail stations. In the calculation of the model, the value of safe value-reduced output is about 680,000 yuan, and the value of safe value-added output is about 970,000 yuan. The value-added benefits generated by the safe investment are greater than the detrimental benefits. From this we can see that the quantitative analysis of the safe and economic benefits of transportation stations has resulted in huge benefits from safe investment.

4 Conclusion

(1) Using the improved C-D production function, an economic loss model for urban rail station accidents was established, and statistical significance was analyzed using the statistical toolbox in MATLAB. The results show that the regression equation in the accidental economic loss model is meaningful and has a high degree.
(2) Based on the functions of "value-added" and "value-reduced" of safe investment, combined with the economic loss model of station accidents, a metropolitan rail transit station safe investment economic model was established and an application example was given. The results show that through the quantitative analysis of the economic benefits of the safe investment, the benefits of safe investment are enormous.
(3) Only safe investment and productive investment with the greatest impact were considered in the establishment of the model. If other investment sub-items have significant effects, further research is needed.

References

1. Wang Y, Li M, Yang B, Yang C (2012) An urban rail transit hazard evaluation methodology based on grey system theory. Procedia Soc Behav Sci 43:764–772
2. Li Y, Wang H, Wang C, Huang Y (2017) Personnel evacuation research of subway transfer station based on fire environment. Procedia Eng 205:431–437
3. Kojić V (2017) Solving the consumer's utility-maximization problem with CES and Cobb-Douglas utility function via mathematical inequalities. Optim Lett 11(4):875–884
4. Jia S, Long Q, Wang RY, Yan J, Kang D (2016) On the inapplicability of the Cobb-Douglas production function for estimating the benefit of water use and the value of water resources. Water Res Manage 30(10):3645–3650
5. Lei D, Chen Z, Deng J (2014) Simplification of production Cobb-Douglas production functions and optimized resource allocation in circular economy of energy and chemical industry. Syst Eng Theory Pract 34(3):683–690
6. Dong J, Huang F (2013) An empirical application based on semiparametric modified Cobb-Douglas production function. Stat Decis 1:92–94
7. Peng H, Li X, Zhang Y, Zhang W (2007) Measuring method of economic benefit of safety investment. Stat Decis 21:163–164
8. Ping X (2003) Exploratory discussion on relationship between input economic benefits of enterprises. World Nonferrous Metals (9):57–58
9. Rune E (2002) How would setting policy priorities according to cost-benefit analysis affect the provision of road safety. Accid Anal Prev 868(12):1–14
10. Peng H, Li X, Li Y (2006) Numerical optimization model of safety investment. Coal Eng 65(12):70–73

Recognition of Fatigue Driving Based on Steering Operation Using Wearable Smart Watch

Dihua Sun, Yong Huang, Min Zhao, Dong Chen and Weijian Han

Abstract Given the growing popularity of wearable smart watch with the capability to detect human hand movements, this paper studies the potential to recognize fatigue driving based on steering operation by using a wearable smart watch. The sensor data used includes acceleration and angular velocity data related to drivers' operation behavior. We analyze the sensors' data features of smart watch under drivers' fatigue and normal states, and select 13 principal characteristic parameters by using the method of principal component analysis (PCA). Then the recognition model of fatigue driving based on support vector machine (SVM) is established. The results show that the proposed method recognizes the drivers' fatigue or normal state more effectively than other methods and its accuracy can reach 83.29%.

Keywords Fatigue driving · Operation behavior · Smart watch · Support vector machine

1 Introduction

Driving is a hard task that drivers always have to watch road and make right operation decisions on time. They are subsequently prone to fatigue after driving for a long time and the ability to control the vehicle will be weakening. Studies show more than 35% of serious traffic accidents are caused by fatigue driving [1]. Therefore, it is necessary to recognize driver's fatigue state effectively by some advanced methods for reducing traffic accidents caused by fatigue driving.

In the existing literature, researchers have made full use of many factors to recognize fatigue driving, which include drivers' facial expression [2, 3], physiological

D. Sun · Y. Huang (✉) · M. Zhao · D. Chen
Key Laboratory of Dependable Service Computing in Cyber Physical Society of Ministry of Education, Chongqing University, Chongqing 400044, China
e-mail: huangyong2016@cqu.edu.cn

College of Automation, Chongqing University, Chongqing 400044, China

W. Han
College of Automotive Engineering, Chongqing University, Chongqing 400044, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_18

characteristics [4, 5] and operation behavior [6–8]. Facial expression-based methods are recognized by analyzing changes of facial expression, such as yawning and blinking. Nevertheless, many external factors, such as illumination and head position change, have an important impact on recognition accuracy. Physiological characteristics-based methods are high in accuracy, but majority of them need to place many sensors on drivers' body for data acquisition, which often make the drivers uncomfortable. Operation behavior-based methods recognize drivers' fatigue state by analyzing operation changes of steering wheel and vehicle status, such as the standard deviation of the steering wheel angle and lane departure, which can avoid the shortcomings of the previous methods and have become an important part of fatigue driving recognition research. Wei et al. [9] chose the lateral position and steering wheel angle as the input data of neural network and trained it by BP algorithm. His experiment results showed that this method is effective to recognize drivers' fatigue level. Mcdonald et al. [10] analyzed the characteristics of drivers' steering data and established a fatigue driving recognition model based on random forest algorithm, which had a higher classification accuracy.

However, the existing majority of operation behavior based methods are to collect the related data by installing various sensors on steering wheel and vehicle, which are too complicated to be practical. In recent years, wearable smart watch with a variety of sensors, have been widely popularized and used to detect human body motion in some research works. Mannini et al. [11] utilized wrist-worn accelerometer to collect wrist and ankle motion data and classified behavior into four classes: sedentariness, cycling, ambulation and other, and the results showed high classification accuracy for ankle data (95.0%) and wrist data (84.7%). Shoaib et al. [12] used wrist-worn motion sensors to recognize hand gestures and it can accurately detect 13 activities. These researches have showed that motion sensors in wearable devices can accurately reflect human's behavior.

According to the above analysis, this paper studies a fatigue driving recognition method based on steering operation by using a smart watch. Thirteen characteristic parameters are extracted from smart watch's sensors based on driving simulation experiment and the fatigue driving recognition model is established based on support vector machine. The test results prove that we can use smart watch to recognize the drivers' fatigue or normal state. This work uses no other external sensors, bringing about a low cost, user-friendly and simple recognition system.

2 Experiments of Data Acquisition

2.1 Experimental Design

Driving experiments are carried out on the Multi-Vehicle Cooperative Simulation Platform, as shown in Fig. 1, in Key Laboratory of Dependable Service Computing in Cyber Physical Society Ministry of Education (CPS-DSC), Chongqing University.

Fig. 1 Driving simulation platform

Fig. 2 Simulation scene

The simulation road is a 100 km long highway of part of the Yuxiang highway located between Chongqing and Hunan province, with two lane in one direction. The smart watch used in the experiment is TicWatch, and it has acceleration sensor, gyroscope sensor, gravity sensor and geomagnetic sensor. Ten participants (six men and four women have an average of 3.5 years driving experience) aged between 24 and 30 years old took part in the experiment from 13:00 pm to 14:30 pm, who are most likely to be fatigued in this period. They have had a good command of the driving scene before the experiment and are asked to wear smart watch for driving during experiments, as shown in Fig. 2. Besides, all participants are required to reduce the lane change and keep the speed of 80 $\pm$ 10 km/h when the vehicle is moving.

2.2 Data Acquisition

In the process of the whole experiments, three types of data are recorded synchronously, including:

(1) The data values of smart watch's acceleration sensor, gyroscope sensor, gravity sensor, and geomagnetic sensor are recorded at the frequency of 10 Hz;
(2) Video of the drivers' facial expression;
(3) The drivers' self-evaluation of fatigue or normal state every ten minutes according to the Karolinska Sleepiness Scale (KSS) [13].

The video of drivers' facial expression and the drivers' self-evaluation are used to judge the drivers' states. There may be a deviation between drivers' self-evaluation and actual fatigue degree. In order to improve the accuracy of the fatigue degree, three trained experts evaluate the drivers' states according to the video in different periods and modify the KSS score of the drivers' self-evaluation. In this study, driving state of the driver is classified into two classes. If KSS score ≤ 3, the driver is under normal state and is under fatigue state while KSS score ≥ 7 [14].

3 Methods

3.1 Method Design

In order to establish the fatigue recognition model and recognize the drivers' states effectively, there are several key steps to do, including driving simulation experiment, coordinate system transformation, characteristic extraction (principal component analysis), training and testing recognition model of fatigue driving based on SVM, as shown in Fig. 3.

3.2 Data Processing

The motion sensors of smart watch use a standard 3-axis device coordinate system to express data values. The X-axis is horizontal and points to the right, the Y-axis

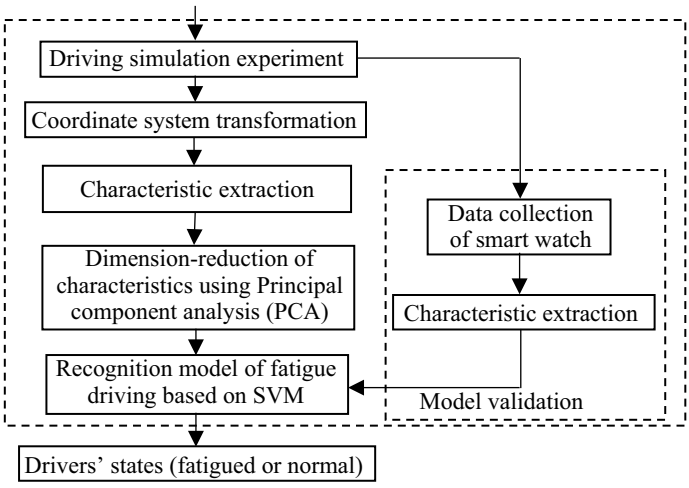


Fig. 3 Flow diagram of fatigue recognition model

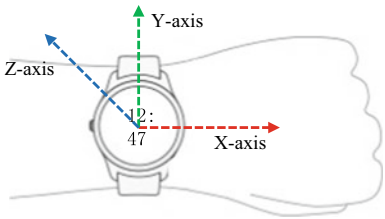


Fig. 4 Device coordinate system

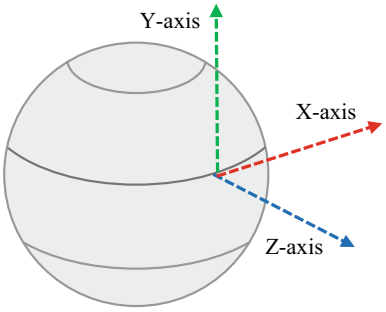


Fig. 5 World's coordinate system

is vertical and points up, and the Z-axis points toward the outside of the screen, as shown in Fig. 4. The attitude of smart watch will be constantly changing in the course of data acquisition and the data values of gravitational acceleration on the three axes will be changing too. Therefore, it is critical to eliminate the effect of gravitational acceleration on the data values of acceleration sensor.

We transform the data values of acceleration sensor from device coordinate system to the world's coordinate system whose X-axis is tangential to the ground and roughly points East, Y-axis is tangential to the ground and points towards the magnetic North Pole, and Z-axis points towards the sky vertically (as shown in Fig. 5):

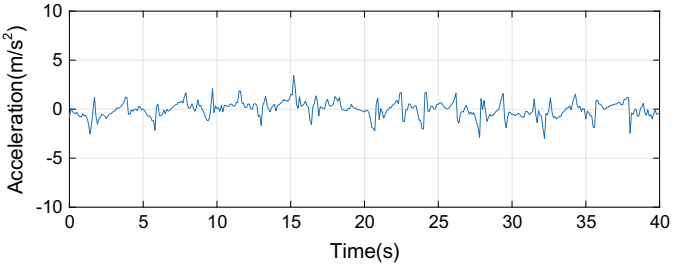
$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} R_1 & R_2 & R_3 \\ R_4 & R_5 & R_6 \\ R_7 & R_8 & R_9 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{1}$$

where x, y, z express the values of acceleration sensor in the device coordinate system and x', y', z' are in the world's coordinate system, R expresses a 3×3 row-major matrix that can be computed with the values of gravity sensor and geomagnetic sensor. When the values of acceleration sensor in the world's coordinate system is got, only the value of Z-axis contains the constant of gravitational acceleration and we can eliminate it directly.

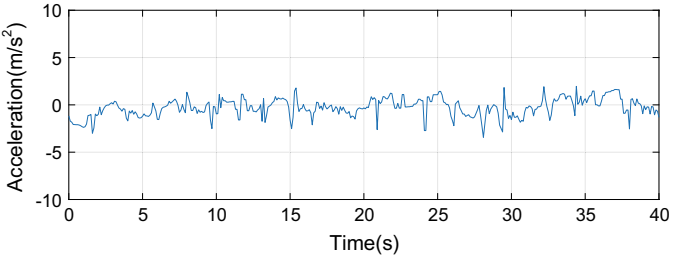
3.3 Characteristics Extraction

3.3.1 Analysis and Extraction of Characteristics

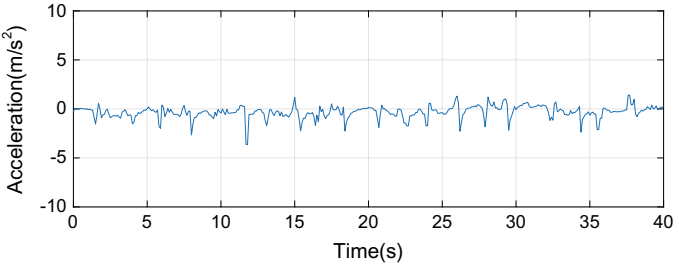
The data of drivers' steering operation includes rotation acceleration and angular velocity of steering wheel that we can get from smart watch's motion sensors. Take rotation acceleration for an example, Figs. 6 and 7 are the data change graphs of rotation acceleration (relative to the world's coordinate system) under normal driving and fatigue driving respectively. It can be seen from the graphs that the steering wheel is corrected frequently and the acceleration amplitude of correction is small under normal driving. However, the correction frequency decreases and the correction



(a) Acceleration of X-axis

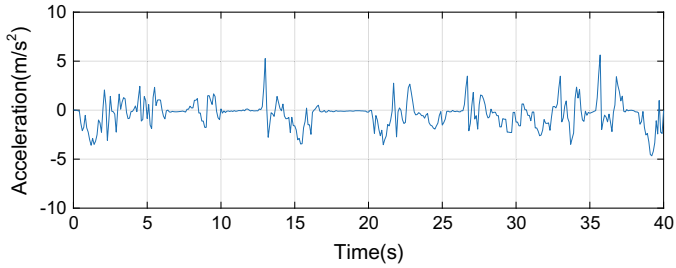


(b) Acceleration of Y-axis

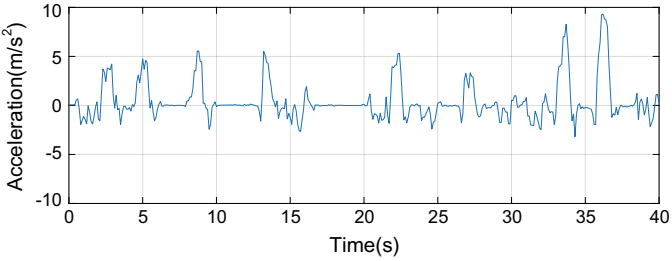


(c) Acceleration of Z-axis

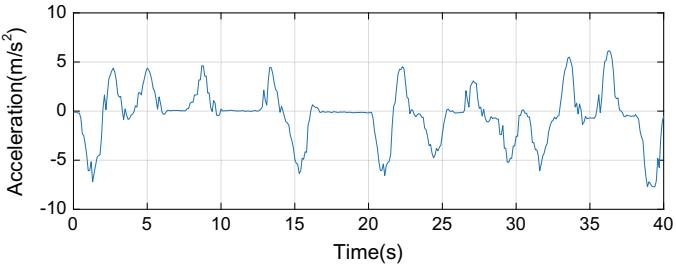
Fig. 6 Acceleration under normal driving



(a) Acceleration of X-axis



(b) Acceleration of Y-axis



(c) Acceleration of Z-axis

Fig. 7 Acceleration under fatigue driving

amplitude of acceleration becomes larger under fatigue driving. To summarize, the correction frequency and amplitude are different under two different driving states.

In order to describe different driving states more accurately, a series of statistical characteristic parameters are defined as listed in Table 1. Taking the duration of fatigue operation and real-time of the recognition model into account, the time window in data series is ten seconds when extracting each characteristic parameter. PLAA and PSAA mean the percentage of the acceleration amplitude that is greater or less than a threshold of the time window:

$$PLAA = \frac{\sum_{i=1}^{n} LA_i}{n} \times 100\%, LA_i = \begin{cases} 1, \ if \ A_i \geq mean\_Q2(AF) \\ 0, \ if \ A_i < mean\_Q2(AF) \end{cases} \quad (2)$$

Table 1 Characteristic parameters and meaning

Characteristic parameters	Meaning
MA	Mean acceleration of steering wheel (X-axis/Y-axis/Z-axis)
SDA	Standard deviation of steering wheel acceleration (X-axis/Y-axis/Z-axis)
RMSA	Root mean square of steering wheel acceleration
PLAA	Percentage of large acceleration amplitude
PSAA	Percentage of small acceleration amplitude
MAV	Mean angular velocity of steering wheel (X-axis/Y-axis/Z-axis)
SDAV	Standard deviation of steering wheel angular velocity (X-axis/Y-axis/Z-axis)
RMSAV	Root mean square of steering wheel angular velocity
PLAVA	Percentage of large angular velocity amplitude
PSAVA	Percentage of small angular velocity amplitude

$$PSAA = \frac{\sum_{i=1}^{n} SA_i}{n} \times 100\%, SA_i = \begin{cases} 1, if\ A_i < mean\_Q2(AN) \\ 0, if\ A_i \geq mean\_Q2(AN) \end{cases} \quad (3)$$

where n is the data number of the time window, A_i expresses the value of synthetic acceleration in data series, $mean\_Q2(AF)$ is the mean value of synthetic acceleration that is greater than median in data series under fatigue state, and $mean\_Q2(AN)$ is the mean value of synthetic acceleration that is greater than median in data series under normal state. PLAVA and PSAVA mean the percentage of the angular velocity amplitude that is greater or less than a threshold of the time window:

$$PLAVA = \frac{\sum_{i=1}^{n} LV_i}{n} \times 100\%, LV_i = \begin{cases} 1, if\ AV_i \geq mean\_Q2(AVF) \\ 0, if\ AV_i < mean\_Q2(AVF) \end{cases} \quad (4)$$

$$PSAVA = \frac{\sum_{i=1}^{n} SV_i}{n} \times 100\%, SV_i = \begin{cases} 1, if\ AV_i < mean\_Q2(AVN) \\ 0, if\ AV_i \geq mean\_Q2(AVN) \end{cases} \quad (5)$$

where n is the data number of the time window, AV_i expresses the value of synthetic angular velocity in data series, $mean\_Q2(AVF)$ is the mean value of synthetic angular velocity that is greater than median in data series under fatigue state, and $mean\_Q2(AVN)$ is the mean value of synthetic angular velocity that is greater than median in data series under normal state. We got a total of 953 samples and randomly selected 600 samples are used to train the recognition model and the other samples are used to test the model.

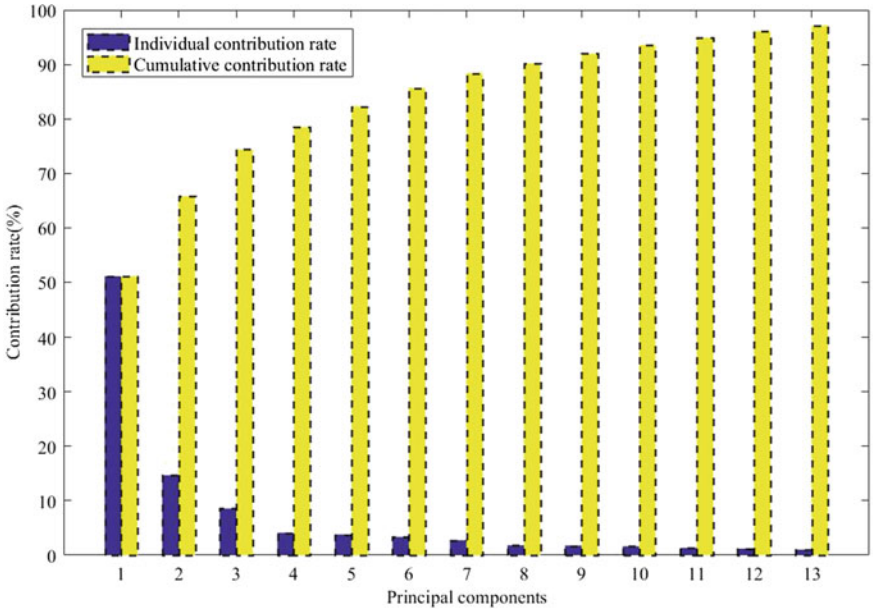


Fig. 8 Contribution rate of the principal components

3.3.2 Principal Component Analysis

We know the number of characteristic parameters is far more than ten because the coordinate system has three axes in Table 1. There may be a strong correlation among these parameters and they have a significant impact on the performance of the recognition algorithm. Thus, it is import to reduce dimensionality of these parameters consisting of some interrelated parameters, while retaining as much as possible of the variation present in them. In this paper, the normalized characteristic data is analyzed by the method of Principal Component Analysis (PCA) and the number of characteristic parameters is reduced from 18 to 13. As shown in Fig. 8, the final 13 principal component parameters are inputs of the recognition model.

3.4 Recognition Model

3.4.1 Support Vector Machine (SVM)

Since the SVM was firstly proposed in 1995 by Cortes and Vapnik [15], it had been widely used for its good classification performance. The principle of support vector machines is to find an optimal hyperplane that separates two classes and is at equal distance from the two. The margin between the optimal hyperplane and the two

classes is maximal. As shown in Fig. 9, H is an optimal hyperplane and it is at equal distance from H_1 and H_2.

Suppose a sample set to be $\{(x_i, y_i)\}, i = 1, 2, \ldots, m, x \in R^n, y_i \in \{-1, 1\}$, where there are m samples for training the SVM, R^n is the n-dimensional feature vector and y_i is the category label. When the feature vectors of samples in different classes are linearly separable, the hyperplane $\omega \cdot x + b = 0$ must satisfy the following inequalities that can separate all samples accurately:

$$y_i(\omega \cdot x_i + b) \geq 1, i = 1, 2, \ldots, m \tag{6}$$

The margin can be calculated as:

$$\min_{\{x_i | y_i = 1\}} \frac{\omega \cdot x_i + b}{\|\omega\|} - \max_{\{x_i | y_i = -1\}} \frac{\omega \cdot x_i + b}{\|\omega\|} = \frac{2}{\|\omega\|} \tag{7}$$

The optimal hyperplane requires the maximization of the margin, which is to mean:

$$\max \frac{1}{2}\|\omega\|^2 = \frac{1}{2}(\omega \cdot \omega) \tag{8}$$

After using the Lagrange dual transformation, the optimal solution of the following objective function is the optimal hyperplane:

$$\begin{cases} \max \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \\ s.t.\ y_i[(\omega \cdot x_i) + b] - 1 \geq 0, \sum_{i=1}^{m} y_i \alpha_i = 0, \alpha_i \geq 0, i = 1, 2, \ldots, m \end{cases} \tag{9}$$

where α_i is Lagrange multiplier of sample data i.

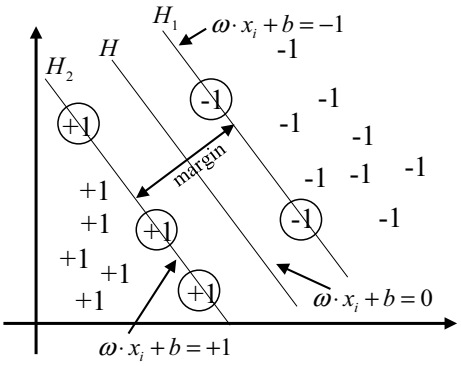


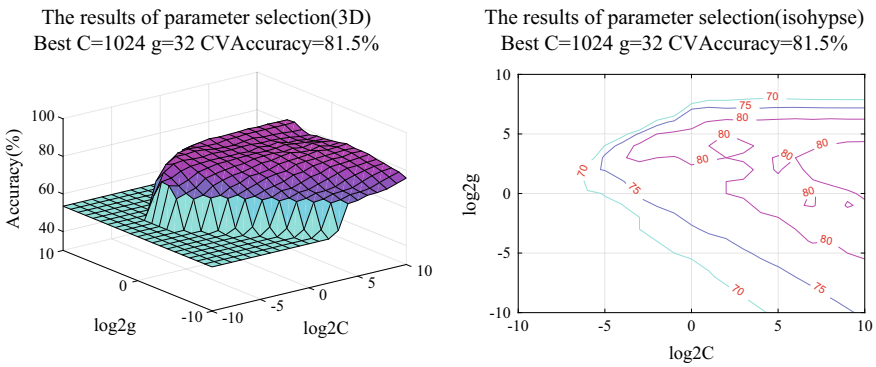
Fig. 9 Optimal hyperplane

When the classification problem is nonlinearly separable, the n-dimensional feature vector can be mapped into a higher dimensional space by using the mapping relationship of $z = \Phi(x)$ and the nonlinear classification problem will be linearly separable by using appropriate kernel function $k(x_i, x_j)$ in high-dimensional space. Then, the optimal solution of the following objective function is the optimal hyperplane:

$$\begin{cases} \max \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j k(x_i \cdot x_j) \\ s.t.\ y_i[(\omega \cdot z_i) + b] - 1 \geq 0, \sum_{i=1}^{m} y_i \alpha_i = 0, \alpha_i \geq 0, i = 1, 2, \ldots, m \end{cases} \tag{10}$$

3.4.2 Model Training

To recognize the driver's fatigue state accurately, this paper uses SVM algorithm for fatigue recognition. Polynomial kernel and radial basis function (RBF) kernel are two popular kernels commonly used in SVM classification at present. Considering that the RBF kernel can realize nonlinear mapping and less hyper parameters can reduce the complexity of the SVM model compared to polynomial kernel, we choose RBF as the kernel for recognition model of fatigue driving. There are two parameters C and g are not known before model training, where penalty factor C expresses the tolerance of error and g in the RBF kernel has an important impact on model prediction. In order to improve the prediction performance of the model, the grid search algorithm is used to optimize the parameters C and g in this paper and the results of optimization are shown in Fig. 10. Finally, the recognition model is trained by using the two optimal parameters and randomly selected 600 samples collected from previous experiments.



(a) 3D view (b) Contour map

Fig. 10 Results of parameter optimization

4 Experimental Results and Discussion

After trained, the recognition model is tested with the other 353 samples, including 168 samples of normal state and 185 samples of fatigue state, and the test results are shown in Table 2. Accuracy is calculated as:

$$accuracy = (TP + TN)/T \times 100\% \tag{11}$$

where TP expresses the number of fatigue state correctly recognized as fatigue; TN expresses the number of normal state correctly recognized as normal, and T means the total test sample number of fatigue and normal state.

Table 2 shows the recognition model, which can classify the drivers' states into two classes, correctly identifies 131 samples of normal state and 163 samples of fatigue state and the accuracy of classification is 83.29%. Figure 11 shows the performance of the used model in this paper and the other related models for fatigue driving recognition. All of these models are trained and tested using the same samples with previous SVM. From Fig. 11, we can conclude that the classification accuracy of SVM used in this paper is higher than other models.

Besides, it is noticed that 37 of 168 normal state samples are incorrectly classified and misjudgment is up to 22.02% in Table 2. The reason for the high misjudgment may be that the current characteristic parameters are not enough and the personal driving style may have an impact on it. In order to improve the recognition accuracy, we will further analyze the data characteristics under different driving states and

Table 2 The test results of the recognition model

The drivers' actual state	The recognition results of model		Total	Accuracy (%)
	Fatigue	Normal		
Fatigue	163	22	185	83.29
Normal	37	131	168	

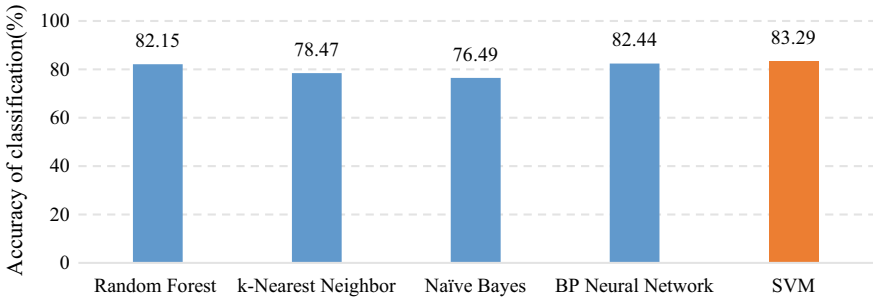


Fig. 11 Performance comparison for different methods

combine with other recognition methods. For the impact of driving style, we will also consider building personalized recognition model of fatigue driving by using smart watch in the following study.

5 Conclusion

The steering operation of a driver under fatigue and normal states will be different. This paper designs and implements a recognition method of fatigue driving based on steering operation by using wearable smart watch. In this proposed method, thirteen principal characteristic parameters are selected from smart watch's sensors by using PCA, and fatigue driving recognition model is established based on SVM. The test results prove that we can use smart watch to recognize the drivers' fatigue or normal state effectively. We also compare the used model in this paper to the other related models. Consequently, the classification accuracy of SVM is better than the other models. This method uses no other external sensors, is low cost, user-friendly and simple. We will further consider more characteristics and build personalized recognition model by using smart watch to improve the accuracy.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 61573075), the National Key R&D Program (Grant No. 2016YFB0100904), the Natural Science Foundation of Chongqing (Grant No. cstc2017jcyjBX0001).

References

1. Khushaba RN, Kodagoda S, Lal S et al (2010) Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. IEEE Trans Biomed Eng 58(1):121–131
2. Filtness AJ, Anund A, Fors C et al (2014) Sleep-related eye symptoms and their potential for identifying driver sleepiness. J Sleep Res 23(5):568–575
3. Mandal B, Li L, Wang GS et al (2017) Towards detection of bus driver fatigue based on robust visual analysis of eye state. IEEE Trans Intell Transp Syst 18(3):545–557
4. Wang H, Zhang C, Shi T et al (2015) Real-time EEG-based detection of fatigue driving danger for accident prediction. Int J Neural Syst 25(2):1550002
5. Chui KT, Tsang KF, Chi HR et al (2016) An accurate ECG-based transportation safety drowsiness detection scheme. IEEE Trans Ind Inf 12(4):1438–1452
6. Chang TH, Hsu CS, Wang C et al (2008) Onboard measurement and warning module for irregular vehicle behavior. IEEE Trans Intell Transp Syst 9(3):501–513
7. Zhang X, Cheng B, Feng R (2010) Real-time detection of driver drowsiness based on steering performance. J Tsinghua Univ 50(7):1072–1076
8. Kaplan S, Guvensan MA, Yavuz AG et al (2015) Driver behavior analysis for safe driving: a survey. IEEE Trans Intell Transp Syst 16(6):3017–3032
9. Wei LI, Chang HQ, Min FX (2010) Detection of driver's fatigue based on vehicle performance output. J Shanghai Jiaotong Univ 44(2):292–296
10. Mcdonald AD, Lee JD, Schwarz C et al (2014) Steering in a random forest: ensemble learning for detecting drowsiness-related lane departures. Hum Factors 56(5):986–998

11. Mannini A, Sabatini AM, Intille SS (2013) Human gait detection from wrist-worn accelerometer data. Gait Posture 37(1):26–27
12. Shoaib M, Bosch S, Incel OD et al (2016) Complex human activity recognition using smartphone and wrist-worn motion sensors. Sensors 16(4):426
13. Akerstedt T, Gillberg M (1990) Subjective and objective sleepiness in the active individual. Int J Neurosci 52(1–2):29–37
14. Chuan XU, Wang X, Chen X et al (2015) Driver drowsiness level analysis and predication based on decision tree. J Tongji Univ 43(1):75–81
15. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

Uncertainty in Lanzhou-Xinjiang Railway Track Longitudinal Level Irregularity Degradation

Ye Yang and Fu-tian Wang

Abstract The state of track deterioration is a key factor affecting traffic safety. In order to better study the deterioration of the track state, it is necessary to analyze firstly the uncertainty of track degradation. This paper analyzes the geometric deterioration of railway track and the uncertainty related to the degradation phenomenon based on a series of data collected from the track sections of K820–K840 railway section of Lanzhou-Xinjiang railway line. Firstly, the degradation parameters of the track irregularity are selected as the research object of this paper, and the evolution trend is studied. The linear relationship between the standard deviation of longitudinal level irregularity and the operation time is proved, and the initial standard deviation and deterioration rate are obtained; then, proper probabilistic distributions are fitted using K-S goodness-of-fit test. The lognormal distribution was selected to establish the relevant deterioration parameter model. The uncertainty of the track geometric degradation was also explored.

Keywords Track degradation · Longitudinal-level irregularity · Uncertainty

1 Introduction

The deterioration state of the track geometry plays a decisive role in the safety of train operations, passenger comfort, equipment life, and track maintenance costs [1]. According to the degradation state of track geometry, the effective management of the track quality in real time, improving the safety of train operation, and prolonging the service life of track equipment, thus reducing the life cycle cost of track equipment have become the research focus of track conservation in various countries. Many experts and scholars at home and abroad have conducted extensive research on the deterioration of the railway infrastructure and its related uncertainties. For instance, Andrade and Teixeira [2] simulated track deterioration and uncertainty associated with degradation phenomenon by using the inspection, operation and maintenance

Y. Yang · F. Wang (✉)
State Key Lab of Rail Traffic Control & Safety, Beijing Jiaotong University, Beijing 100044, China
e-mail: 16120917@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_19

data collected on the upgrade line from Lisbon-Porto in Portugal. Andrade and Teixeira [3] proposed a Bayesian model to assess rail track geometry degradation, building up a framework to update the uncertainty in rail track geometry degradation throughout its life cycle.

In this paper, a series of vehicle inspection data collected from the K820 to K840 sections of the Lanzhou-Xinjiang railway line is used as a reliable data to study the degradation of the track. At first, the degradation parameters of the track are selected as the research object of this paper, and its evolution trend is studied. Then the lognormal distribution is selected to establish the corresponding degradation parameter model, and the uncertainty of the degradation parameter is studied.

2 Track Geometry Degradation Parameters

2.1 Selection of Degradation Parameters

Track geometry degradation is generally studied by evaluating the deterioration of geometric parameters. In general, the inspection indicators of the track inspection vehicle include longitudinal level irregularity, gauge irregularity, horizontal-alignment irregularity, cross-level irregularity and track twist irregularity, etc. In fact, many infrastructure managers tend to focus all these defect types into the track quality index. The track quality index (TQI) [4] is an index for assessing the track state proposed by the United States Railway Corporation, which is a comprehensive index and evaluation method that uses mathematical statistical methods to describe the overall quality status of the section track. The TQI is the sum of the standard deviations of the irregularity management indicators, such as longitudinal level irregularity, gauge irregularity, horizontal-alignment irregularity, cross-level irregularity and track twist irregularity in a 200 m track section.

However, as confirmed by the International Union of Railways (UIC) [5], for short waves, the standard deviation is still a key parameter for maintenance strategies. A recent study [6] founds that the trend of deterioration of the track longitudinal-leveling after maintenance actions can be well fitted with a linear relationship, and the slope of the fitting line is close to each time after tamping, indicating the deterioration rate of longitudinal-leveling remains unchanged; the deterioration rate of longitudinal level irregularity is much greater than the track twist, gauge deviations. Therefore, this article only analysis the uncertainty associated with longitudinal level irregularity. It is defined as the geometrical error in the vertical plane, which is represented by the difference between a point on the top of the rail in the running plane and the ideal average of the longitudinal profile.

In the past, many experimental studies have confirmed the linear relationship between the longitudinal-leveling standard deviations and the operation time. The China railway is generally checked once a month with a track inspection vehicle,

and a large amount of track inspection vehicle data is accumulated, which is conducive to the analysis of the deterioration trend of the rail line status. Therefore, it is more convenient to study the relationship between the longitudinal-leveling standard deviations and time.

In order to make better use of data information for research, track inspection records should be condensed, usually considering a 200 m long track section.

Use the following linear relationship to estimate the evolution of the standard deviation for each 200 m-segment:

$$\sigma = c_1 + c_0 T$$

in which σ is standard deviation of longitudinal level irregularity (mm); c_1 is initial standard deviation (mm); c_0 is deterioration rate (mm/month); T is time cycle (one month).

2.2 *Study on the Evolution Trend of Deteriorating Parameters*

This study randomly selected 12 sets of the standard deviation data of longitudinal level irregularity from K800 to K840 and April 2015 to April 2016 of the Lanzhou-Xinjiang railway line for linear fitting, Fig. 1 shows the results of the fitting. Note that in Fig. 1, though the linear fitting is better, the deterioration parameters are relatively different, which means that the initial standard deviation and the deterioration rate of each track segment are different.

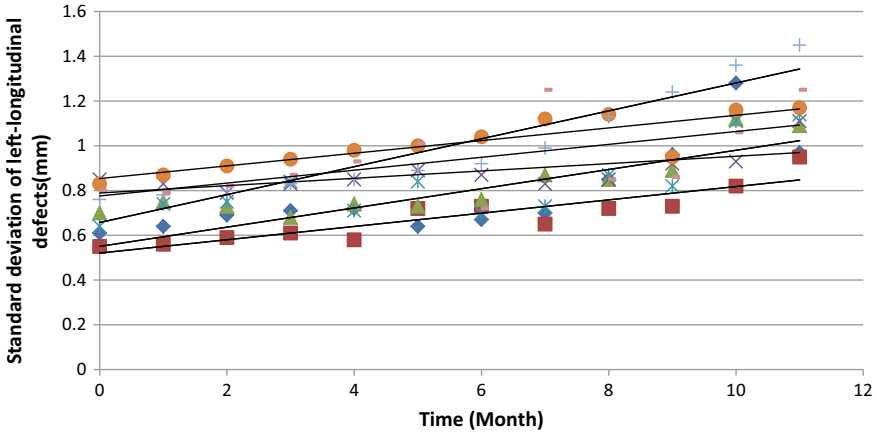


Fig. 1 Example of the evolution of the standard deviation of longitudinal level irregularity for 200 m track sections

Table 1 Significant levels of linear fit for the standard deviations of longitudinal level irregularity for different track sections

Group no.	Sections	R^2	Group no.	Sections	R^2
1	800.4	0.8094	5	824.2	0.5754
2	808.8	0.6188	6	828.6	0.7654
3	811.2	0.7426	7	833.6	0.8937
4	816.8	0.5049	8	837.8	0.353

The significance level of the linear fit is relatively high for the data of standard deviation deterioration trends of longitudinal level irregularity, which randomly selected from 200 m track sections, indicating it is a good liner relationship between the standard deviation of longitudinal level irregularity and the operating time Table 1 shows the significance levels of fitting.

3 Degradation Parameters Uncertainty

3.1 Selection of Deteriorating Parameter Model

Studies [7] carried out by the Office for Research and Experiments in the 1970s and recent research [6] show that the deterioration rates of different sections are different, but are still a constant parameter throughout the life cycle. In order to assess the uncertainty associated with the track geometry degradation, coefficients c_1 and c_0 are defined as the random variable, and typical probability distributions were tried, such as normal distribution, lognormal distribution, uniform distribution, exponential distribution, and Poisson distribution. The distributions are defined as follows.

(1) The normal distribution is defined by its probability density function and respective parameters:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

in which μ is scale parameter and σ is shape parameter.

(2) The lognormal distribution is defined by its probability density function and respective parameters:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}[\ln(x)-\ln(\mu)]^2}$$

in which μ is scale parameter and σ is shape parameter.

(3) The uniform distribution is defined by its probability density function and respective parameters:

$$f(x) = \frac{1}{b-a}, a \le x \le b$$

in which a is minimum parameter and b is maximum parameter.

(4) The exponential distribution is defined by its probability density function and respective parameters:

$$f(x) = \lambda e^{-\lambda x}$$

in which λ is rate parameter.

(5) The poisson distribution is defined by its distribution law:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1 \ldots$$

in which λ is the average incidence rate of random events within a unit time (or unit area).

The 100 track units data of deterioration rate and initial standard deviation from K800 to K840 from April 2015 to April 2016 of the Lanzhou-Xining Line were selected, and K-S goodness-of-fit test was performed on the above-mentioned types of distributions. The results show in Table 2.

The results of K-S goodness-of-fit test for all types of distributions show that the Poisson distribution is suitable for discrete data and is not suitable for the modeling of this case. Through the comparative analysis, only the lognormal distribution proved to be suitable for modeling of the Lanzhou-Xinjiang railway line, as the goodness-of-fit Sig. is greater than 0.05, and other goodness-of-fit Sig. is relatively low, such as the normal distribution, the uniform distribution, the exponential distribution. Therefore, the author believes that it is appropriate to model the lognormal variable.

Table 2 Goodness-of-fit tests to various distributions

Distribution types	K-S test sig.
Normal distribution	0.003
Lognormal distribution	0.105
Uniform distribution	0.000
Exponential distribution	0.002

Table 3 Goodness-of-fit tests to lognormal distributions for deterioration rates and initial deviations for first group track sections

	Deterioration rate, c_0 (mm/month)	Initial standard deviation, c_1 (mm)
Scale parameter	1.144	0.623
Shape parameter	0.835	0.325
N	200	200
K-S test sig.	0.200	0.315

Table 4 Goodness-of-fit tests to lognormal distributions for deterioration rates and initial deviations for second group track sections

	Deterioration rate, c_0 (mm/month)	Initial standard deviation, c_1 (mm)
Scale parameter	0.946	0.571
Shape parameter	1.003	0.279
N	200	200
K-S test sig.	0.265	0.546

3.2 Study on Uncertainty of Deterioration Parameters

This study selected randomly 400 track units data of the deterioration rate and initial standard deviation collected from the K820 to K840 sections of rail inspection vehicles of Lanzhou-Xinjiang railway line from April 2015 to April 2016, and divided into two groups to fit using lognormal distributions. The relevant probability distributions are fitted using K-S goodness-of-fit test. The results show in Tables 3 and 4.

As Tables 3 and 4 show that the hypothesis that the deterioration rate and initial standard deviation also follow a lognormal distribution are accepted at a 5% significance level in two sets of data, but their shape parameters and scale parameters are different, so the degradation parameters are uncertain for different segments.

4 Conclusions

This paper analyzed and calculated the degradation rate and initial standard deviation of track longitudinal level irregularity for each track segment based on the data of vehicle inspection vehicles of Lanzhou-Xinjiang Railway. The lognormal distributions are modeling to fit track geometry degradation parameters, and K-S goodness-of-fit tests are performed on parameters. The degradation of track longitudinal level irregularity and the uncertainty associated with the degradation phenomenon are also studied, which has guiding significance for the study of track geometry degradation.

Acknowledgements This work was financially supported by the National Natural Science Foundation of China (grant no. 51578057) and supported by China Railway Corporation Research Project

(grant contract no. 2017T003-C). The writers would like to thank the support of them. Moreover, the writers wish to thank the reviewers for their comments.

References

1. Zhou BH et al (2004) China railway encyclopedia. China Railway Press, Beijing
2. Andrade AR, Teixeira PF (2000) Uncertainty in rail track geometry degradation: Lisbon-Oporto line case study. J Transp Eng 137:193–200
3. Andrade AR, Teixeira PF (2012) A bayesian model to assess rail track geometry degradation through its life-cycle. Res Transp Econ 36(1):1–8
4. Xu P (2009) Study on the characteristics of dynamic irregularity in railway tracks. Beijing Jiaotong University
5. International Union of Railways (UIC) (2008) Best practice guide for optimum track geometry durability. UIC, Paris
6. Yang F (2017) Research on maintenance standards and decision-making technology based on longitudinal-leveling irregularity. Railw Eng 7:131–135
7. Esveld C (2001) Modern railway track, 2nd edn. MRT Productions, Zaltbommel

Functional Areas Layout in Logistics Park Combining Traffic Organization by Genetic Algorithm and Fuzzy Clustering

Qin Xiang

Abstract The functional areas layout and the organization of the traffic will influence with each other in logistics park. In this paper, the author studies the method to combine the layout of the functional areas with organization of the traffic so as to make the layout more practical. Firstly, the model for functional areas layout design is established, which is based on traffic organization. Secondly, the solution algorithm is proposed, which includes three parts. The first part is fuzzy clustering analysis which is used to cluster functional areas so as to determine the traffic network factors. The second part is genetic algorithm for layout problem at the functional areas, the author uses three parts of codes to construct the chromosome encoding, which are code for the area, cutting code and sequence code. The third part is traffic organization, compared with genetic algorithm and CPLEX, the author selects shortest path distance for calculation to speed the calculation speed. And CPLEX would be used to calculate the actual distance after the final layout is obtained. Finally, the model and algorithm are applied in the Lianyungang Port Logistics Park to verify the feasibility. Compared with the base layout, the truck distance can be reduced by 13.8%.

Keywords Functional areas layout · Genetic algorithm · Logistics park · Traffic organization

1 Introduction

In recent years, logistics has become an important factor in China's economic development and a new pillar industry in many cities because of the rapid development of the logistics industry. According to the investigation by China Federation of logistics and purchasing in 2015, there were about 754 domestic logistics parks in 2012 and the number increases to 1210 in 2015 and this significant increment (60%) indicates the major role of the logistics park in Chinese logistics and business development.

Q. Xiang (✉)
Business School, Shanghai Normal University Tianhua College, Shanghai 201815, China
e-mail: 939606613@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_20

Although the scale of logistics park is very large, the vacancy rate is also very high, almost 60% of logistics parks are not in business because of the problems in layout and traffic organization. It is important to connect layout of different functions with the organization of the traffic.

The layout design methods are based on the facility layout problem (FLP), which is concerned with positioning functional areas to locations within a logistics park. Unlike facility layout problem, the functional areas layout design needs to consider the feature of logistics park. For example, the aspect ratio is always used to restrict the occurrence of an extremely long and narrow department in facility layout problem, but this constraint is not reasonable in logistics park layout design. In facility layout problem, the distance between functional areas is Manhattan distance or Euclidean distance, however, this method is not applicable in logistics park layout design since the operation of trucks depend on network of the traffic, so the distance is acquired based on traffic organization. These features complicate the functional areas layout design in logistics park. Therefore, the aim of this study is to demonstrate how to solve the functional areas layout design problem in logistics park combining traffic organization. The contributions of this paper are:

- A new method that transfers from layout to the network is proposed. Based on this method, the traffic network factors can be determined and the actual distance can be calculated, in this way, the layout method can prevent us from restricting the shape with the aspect ratio.
- In order to make the distance more practical, we put forward the traffic organization method on the network of the traffic in logistics park, which makes the distance we use more reasonable.

2 Literature Review

2.1 Facility Layout Problem

Until now, there are lots of researches about the facility layout problem, according to the survey by Drira et al. [1], it can be divided into three part.

The first part involves algorithms addressing the general FLP. For the large scale FLP, there are some heuristic algorithms utilized in this field, such as Tabu search [2], ant colony optimization [3] and so on. The second, some dynamic facility layout problem [4] and multi-objective facility layout problem [5] are covered in the extension of the general FLP, and there are so many algorithms as well. The third, single-row facility layout problem [6] is an example of the structured instances problem, for example, single-row facility layout problem [6].

2.2 Functional Areas Layout Problem in Logistics Park

The FLP has been proven to be NP-complete. Given that the functional areas layout problem in logistics park is at least as difficult as the FLP, it also belongs to the class of NP-complete problems.

In 1961, using the method of systems engineering and systems analysis, Richard Muther put forward a representative method for system layout design—SLP (Systematic Layout Planning), its main idea is to configure each work unit according to the area proportion and relationship graph [7]. Based on this method, many scholars develop new algorithms to solve the layout problem. The research method also can be divided into three parts. The first part is SLP combining heuristic algorithm [8], the relativity between areas can be obtained by SLP, then the heuristic algorithm is used to optimize the layout form, the objective function they considered is to achieve the maximum total compositive relativity. The second part is SLP combining Fuzzy Cluster Analysis [9], based on this method, the reasonable correlation is achieved between the various functional areas of logistics parks. The third part is graph theory [10], the center point and median point of graph theory can be used to solve the location problem of maximum distance to a minimum.

Most of the above approaches to functional areas layout problem in logistics park are based on the SLP, for the SLP, it needs to calculate the relativity between the areas. However, it is very hard to define that. Therefore, many authors choose the total distance of truck to evaluate the layout performance, because during the operation, if two areas have strong relationship, the freight volume would be higher and they should be arranged nearby. In this way, another problem can be raised, that is how to calculate the distance between functional areas. Normally, we can use Manhattan distance or Euclidean distance, but it is not applicable since the traffic need be organized in logistics park based on the traffic network. Therefore, the relationship between the traffic organization and layout of the functional areas need to be considered. Zheng [11] proposed a framework for logistics park planning based on the relation of them, this framework can adjust the layout of functional areas.

From the analysis above, it shows that few studies analyse the relationship between the traffic organization and layout of the functional areas in logistics park so that the actual distance is replaced by Manhattan distance or Euclidean distance. What is more, the related researches are mainly based on qualitative analysis at present, they lack quantitative method. So as for this problem, this paper seeks to design the layout of the logistics park by applying a algorithm based on the relationship between them.

3 The Layout Model of the Functional Areas

For the convenience, this paper presents the network of transportation through a directed graph, $G = (V, A)$. V represents the vertices of the graph, A is the edges of it. I and j are the serial number of the functional areas. N_{ij} stands for the volume of

traffic between i and j. C_a is the capacity of road a. M_a is the distance of road a. The decision variable is x_{ij}^a, it decides whether N_{ij} selects road a or not. So the objective function is shows as following, which is total truck distance.

$$\min Z = \sum_{a \in A} \sum_{i,j \in V} M_a N_{ij} x_{ij}^a \tag{1}$$

In the process of layout, the planners need to make sure the functional areas cannot overlap with each other, so they need to satisfy the constrains as follow:

$$\left|h_i - h_j\right| \geq \frac{l_i + l_j}{2}, \quad i, j = 1, 2, \ldots, n \tag{2}$$

$$\left|z_i - z_j\right| \geq \frac{b_i + b_j}{2}, \quad i, j = 1, 2, \ldots, n \tag{3}$$

In which, (h_i, z_i) is the coordinates of center of functional area i, l_i is the length of functional area i, b_i is the width of functional area i, n is the number of functional areas.

Meanwhile, the planning area includes the functional areas totally, so the constrains are shown as follow:

$$\frac{l_i}{2} \leq h_i \leq L - \frac{l_i}{2}, \quad i = 1, 2, \ldots, n \tag{4}$$

$$\frac{b_i}{2} \leq z_i \leq W - \frac{b_i}{2}, \quad i = 1, 2, \ldots, n \tag{5}$$

In which, L is the length of planning area, W is the width of planning area.

The network of the transportation can be shaped after finishing the layout design. Based on this traffic network, the distance for the trucks can be calculated by organizing the traffic flow.

In order to ensure the management and control of traffic flow, the author decides not to separate each traffic flow. The means each OD (Original-Destination) only chooses one route. So the constrains are shown as follow:

$$\sum_{a \in O_i} N_{ij} x_{ij}^a = N_{ij}, \quad i, j = 1, 2, \ldots, n \tag{6}$$

$$\sum_{a \in D_j} N_{ij} x_{ij}^a = N_{ij}, \quad i, j = 1, 2, \ldots, n \tag{7}$$

In which, O_i is the set of roads whose original is node i, D_i is the set of roads whose destination is node i.

In order to guarantee the equilibrium of traffic in intermediate points, request:

$$\sum_{a \in D_k} x_{ij}^a \leq 1, \quad \forall k \in V, \quad k \neq i, \quad k \neq j \tag{8}$$

$$\sum_{a \in O_k} x_{ij}^a \leq 1, \quad \forall k \in V, \quad k \neq i, \quad k \neq j \tag{9}$$

$$\sum_{a \in D_k} x_{ij}^a - \sum_{a \in O_k} x_{ij}^a = 0, \quad \forall k \in V, \quad k \neq i, \quad k \neq j \tag{10}$$

In order to ensure the total traffic is constrained by road capacity, and the traffic flow distribution as even as possible to prevent certain sections of the flow is too concentrated, requirements:

$$\sum_{i=1}^{V} \sum_{j=1}^{V} N_{ij} x_{ij}^a \leq 0.75 C_a, \quad \forall a \in A \tag{11}$$

4 Algorithm of the Model

There are 3 steps to design the layout of the functional area: firstly, fuzzy clustering is applied to divide the functional areas into clusters so as to determine the traffic network factors; secondly, their sizes and freight volume would decide the layout, then based on this layout and the result of fuzzy clustering, the parameters of traffic network in logistics park can be generated; Thirdly, the method can acquire the plan of traffic organization according to the traffic network, then the plan will be transmitted to the stage 2, which also influences the layout of the functional area.

4.1 Fuzzy Clustering for Functional Areas

During the operation, if two areas have strong relationship, the freight volume would be bigger, so in order to minimize the cost, they should be arranged nearby. In this paper, the author selects fuzzy clustering to divide the functional areas into clusters, the freight volume between functional areas is set as main affecting factor, the factor value range is l to 100, the step size is 10, the minimum value is l, and the value is proportional to the degree of the index, then the initial-data matrix can be set up. However, different data are usually with different dimensions, the data obtained by the method above are not always within the interval [0, 1], so appropriate transformation should normally be done for these data to make them can be compared. The data standardization is to put data within the interval [0, 1] according to requirements of fuzzy clustering matrix. Commonly, the "translation—range transformation" [12] method is used to eliminate the influence of dimensions.

Then similar matrix R needs to be set up. Similar parameter $r_{ij}(r_{ij} \in R)$ is the similar degree of functional area i to functional area j. The methods to determine similar parameter in this paper is Euclidean Distance [12]. After that, the clustering effect can be analyzed, according to the fuzzy clustering tree, the clustering result can be acquired.

4.2 Functional Areas Layout Method

The functional areas layout has been proven to be NP-complete problem, no computationally efficient approach has been found to obtain an optimal solution to the problem, so in this paper, the author selects genetic algorithm [13] to solve it.

The author applies chromosome encoding in the genetic algorithm to ensure the effectiveness and efficiency of it. Chromosome encoding combined with binary tree can avoid the generation of infeasible solutions and improve operation ability in the process of implementation. According to this idea, three kinds of codes are chosen to use in chromosome encoding, which are code for areas, cutting code and cutting sequence code. The first code, code for areas stands for the relative position of different areas. The cutting code can be applied to identify the cutting way and find relationship of different areas. Vertical cutting is marked by 1 and horizontal cutting by 0. The order of cutting is represented by code for the cutting sequence. Meanwhile, as the fuzzy clustering in the first stage has divided the functional areas into clusters, for the cutting sequence code, it needs to divide the planning area into several modules firstly according to the fuzzy clustering result, then the functional areas can be arranged in each module. A logistics park consisted of 5 functional areas is taken for example, it can be divided into 2 clusters, one are functional area 1 and functional area 5, the others are functional area 2, functional area 3 and functional area 4, so the feasible chromosome encoding is shown as in Fig. 1.

During the chromosome decoding, the cutting code needs to be inserted into the module code according to the corresponding sequence, its order depends on the cutting sequence code. In this way, the plane cutting process is completed. The specific steps are as follows:

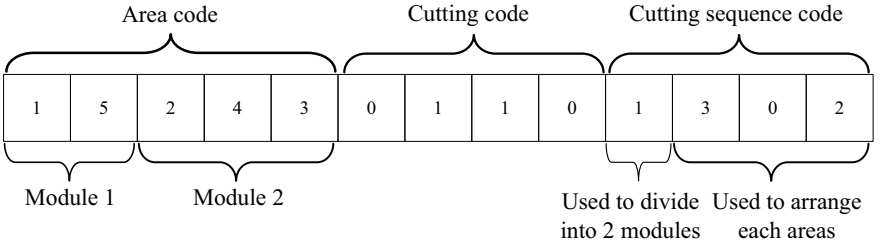


Fig. 1 Constitution of chromosome encoding

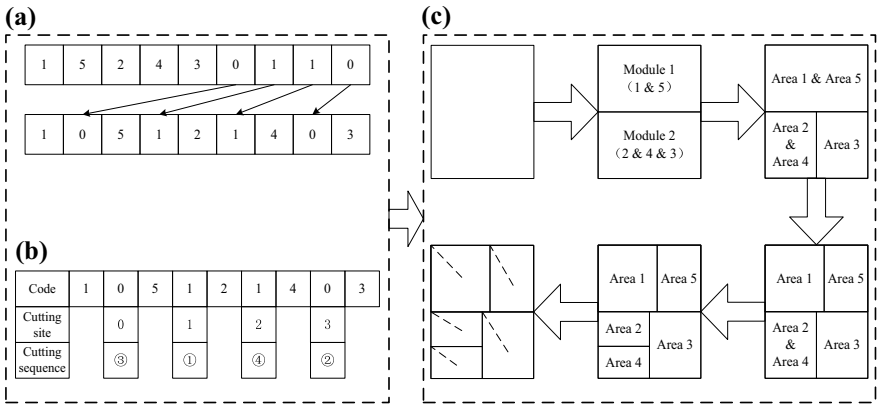


Fig. 2 Chromosome decoding

(1) Chromosome encoding transformation (as shown in Fig. 2a).
(2) Cutting order decision (as shown in Fig. 2b).
(3) Layout transformation and Network generation (as shown in Fig. 2c). When we generate the network, we assume that traffic flow would be produced by the cell centroid, and it can enter into the network by the node in upper left corner.
(4) Network factors decision. According to the result of fuzzy clustering, it shows that the network in logistics park can be separated to two levels. For the road between each module, they need to bear the transportation of goods between modules, so the capacity of these roads should be bigger; for the road in each module, they only need to provide the service for the transportation of goods between functional areas in the modules, so the capacity of these road should be smaller.

In this paper, stochastic tournament model is selected as the selection operator, one-point crossover is selected as the crossover operator, and simple mutation is selected as the mutation operator. What's more, as the chromosome encoding in this paper has three sections, the crossover operator and mutation operator should be applied to each section, and in this process, the chromosome encoding would get infeasible solutions, so they need to be adjusted.

4.3 The Organization the Traffic of the Park

From the analysis above, it shows that the traffic organization result will be transmitted to the second stage to influence the layout of the functional area. When combining the organization of the traffic with the problem of the layout, it is found that the layout of the functional areas should consider the calculation of the traffic organization

plan. Therefore, problem solving speed will depend on the speed of the algorithm for traffic organization.

In the previous study, genetic algorithm was used to solve the problem [14]. It costs about 2 min to solve the traffic organization issues having 12 functional area in some logistics park by the method. It is estimated that the software will take over 266 h to calculate functional layout problem which has 80 generations and 100 populations for one generation by the genetic algorithm.

Currently IBM ILOG CPLEX Optimization Studio (often informally referred to simply as CPLEX) becomes more and more popular, not only for business, but also for research. It can solve the integer programming problems and very large linear programming problems quickly with excellent results. So the author also tries to solve the same traffic organization problem by CPLEX. Fortunately, the time consumption is around 15 s. Therefore, if the scale is as large as the problem we mention above, the software will cost around 34 h. Although calculation speed has been greatly improved, it is still too long.

Hence, the article utilizes the shortest path distance as the actual distance to speed the calculation. The shortest path distance can be got through Dijkstra algorithm easily. Although this method might not find the optimal solution, it can help in reducing the complexity of the problem and can obtain some satisfactory solutions quickly. When we obtain the final layout, we can use CPLEX to calculate the actual traffic organization plan according to the capacity of road.

4.4 Flow Chart

Figure 3 is the flow chart to the model solution.

5 Case Study

The article takes the Lianyungang Port Logistics Park [15] as an example by applying the mathematical algorithm and model in the article.

Functional areas in the park are shown in Table 1.

The volume for the freight traffic is shown in Table 2.

Firstly, SPSS is used for fuzzy clustering, through the analysis, the Fuzzy clustering tree is given in Fig. 4. It shows that the functional areas can be divided into three clusters, the first one is functional area 1, 2, 3 and 4, the second one are functional area 5 and 11, others are the third one. So in the process of network factors decision, according to the freight volume in this park, the capacity of road between each modules is 1000 pcu/h, and the capacity of road in each modules is 600 pcu/h.

The article designs the layout of the park through the software, which is shown in Fig. 5. It takes 1 h and 28 min to finish the calculation. The results keep consistent until the 23th generation. The objective function's value is 65,415 km/h, in which

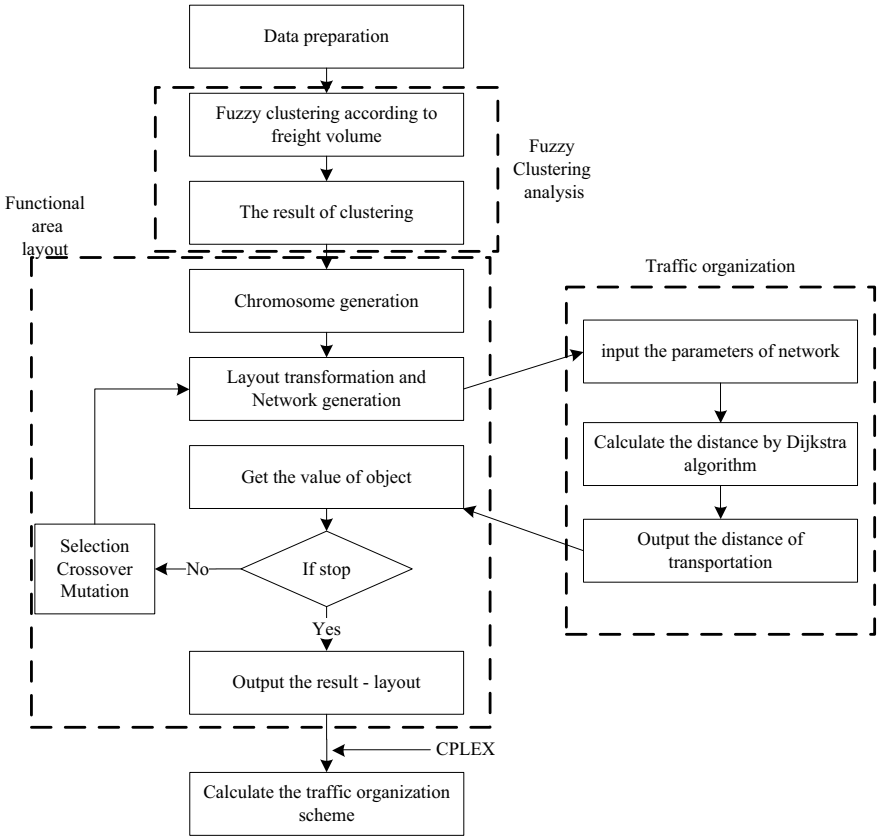


Fig. 3 Flow chart to model solution

Table 1 Functional areas

Item	Area name	Size (ha)	Item	Area name	Size (ha)
1	Container yard	69	7	Comprehensive management area	7
2	Bonded warehouse	58	8	Business district	47
3	General warehouse	93	9	City distribution area	20
4	Bulk yard	51	10	Multimodal transport area	25
5	Industrial area	51	11	Railway facelift area	35
6	Characteristic processing zone	21	12	Life living area	9

Table 2 The freight volume between the functional areas (pcu/h)

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	24	4	3	6	3	0	6	5
2	0	0	0	0	18	43	6	5	3	4	35	11
3	0	0	0	0	62	10	13	15	8	23	134	26
4	0	0	0	0	95	16	29	23	11	25	393	57
5	21	26	54	83	0	0	22	19	7	20	208	44
6	5	13	13	19	0	0	4	3	1	3	21	8
7	4	4	10	16	18	6	0	5	3	6	46	0
8	8	0	18	29	21	5	5	0	3	2	14	10
9	3	0	6	10	7	3	3	3	0	1	9	5
10	0	16	17	78	45	8	14	9	3	0	90	27
11	25	25	87	82	71	12	18	15	5	24	0	35
12	7	8	20	31	35	11	0	10	5	11	91	0

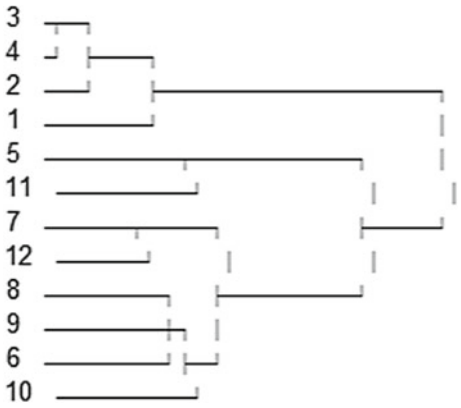


Fig. 4 Fuzzy clustering tree

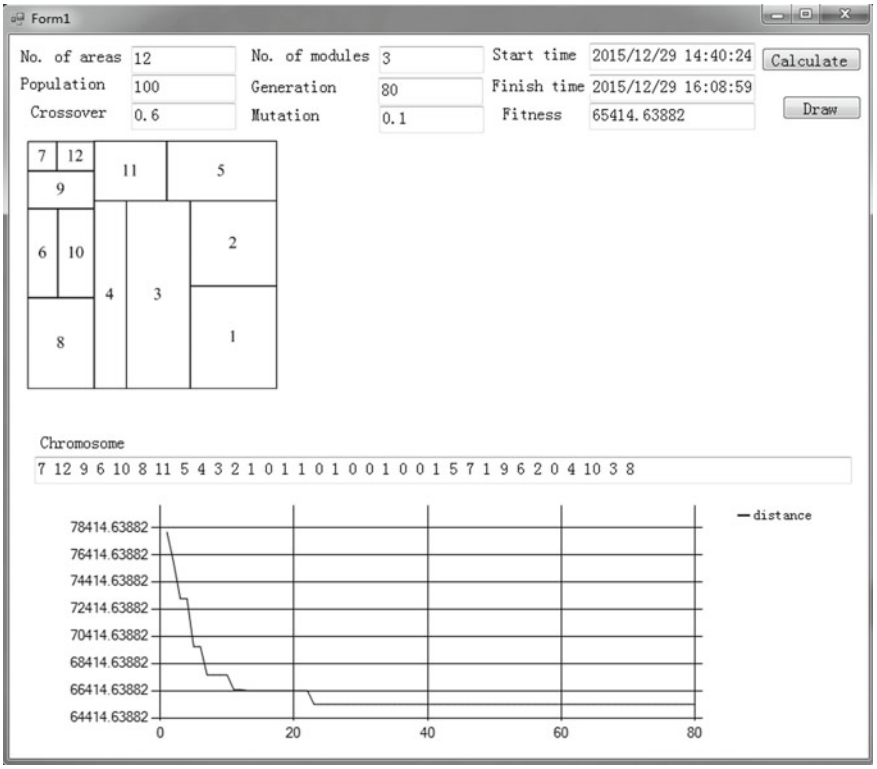


Fig. 5 Logistics park layout based on traffic organization

distance is shortest path distance. Then using CPLEX, the actual traffic organization distance can be obtained, its value is 65,447 km/h. At the same time, the result is given in Table 3 by the comparison. Form the layout, it shows that no extreme situation such as long and narrow functional area shows up. Meanwhile, the distances between the interactions are reasonable. The total distance for the truck can be saved by 13.8% at the most.

In order to verify the stability of the algorithm we proposed, we used the software to solve it for five times, the result is shown in Table 4. From the table, we can know the best value is 64,135 km/h, the worst value is 65,440 km/h, and the gap between them is only 2%. What's more, in the fifth time, all the solution is better than the base result in reference [15]. And sometimes, the objective function value equals to the actual distance value, which means the layout and traffic organization can match each other well so that all the trucks can travel by shortest path.

6 Conclusion

The layout of functional areas and internal traffic in the logistics can interact and influence with each other during the planning of the park. In this paper, the author studies the method to combine the two factors together. First, the model for layout problem of functional areas is established, then the algorithm for the model is also proposed. Through the fuzzy clustering analysis, function areas that have strong relationship are clustered into a large module, and in order to obtain the solution, the author selects genetic algorithm, in which chromosome encoding consists of three sections. The shortest path distance is selected to speed the calculation. Using the algorithm above, the software is developed. Finally, the paper selects the Lianyungang Port Logistics Park as an example. The total distance for the trucks can be saved by 13.8% at the most.

Table 3 Result comparison

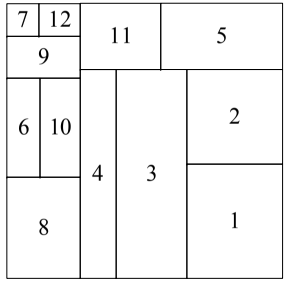
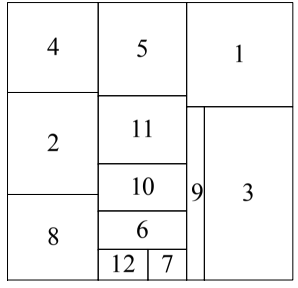
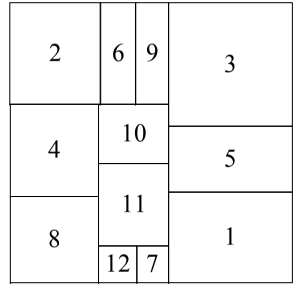
Layout method	Layout combining with traffic organization	SLP [15]	Genetic algorithm based on SLP [15]
Layout form			
Actual distance (km/h)	65,447	74,468	71,200
Distance saving (km/h)	–	9021	5753
Saving percentage (%)	–	13.8	8.8

Table 4 Result comparison for different methods

Time	1	2	3	4	5
Objective function (km/h)	65,164	64,135	64,348	65,128	65,440
Actual distance (km/h)	65,250	64,880	64,348	65,128	65,440

References

1. Drira A, Pierreval H, Hajri-Gabouj S (2007) Facility layout problems: a survey. Annu Rev Control 31(2):255–267
2. Kulturel-Konak S (2012) A linear programming embedded probabilistic tabu search for the unequal-area facility layout problem with flexible bays. Eur J Oper Res 223(3):614–625
3. Wong KY (2010) Solving facility layout problems using flexible bay structure representation and ant system algorithm. Expert Syst Appl 37(7):5523–5527
4. Pourvaziri H, Naderi B (2014) A hybrid multi-population genetic algorithm for the dynamic facility layout problem. Appl Soft Comput 24:457–469
5. Matai R (2015) Solving multi objective facility layout problem by modified simulated annealing. Appl Math Comput 261:302–311
6. Kothari R, Ghosh D (2014) An efficient genetic algorithm for single row facility layout. Optim Lett 8(2):679–690
7. Muther R (1961) Systematic layout planning. Industrial Education Institute, Boston
8. Sun Y, Ma C, Zhang W (2015) Research on layout problem in logistics park based on shortest path and multi-population genetic algorithm. Logist Sci-Tech 2:86–91
9. Zhu X, Zhang Q (2013) A novel method on layout planning for internal functional areas of logistic parks. LISS 2012. Springer, Berlin
10. Jiang C, Bai L, Zheng W (2010) Research on layout of airport logistics park based on graph theory: an empirical study of Ningbo airport logistics park. In: 2010 international conference on intelligent computation technology and automation, Changsha
11. Zheng H (2010) Planning of traffic system and research of the relation with working-area on logistics park. Southeast University, Nanjing
12. He Q, Xue X (2009) A research on layout planning of regional logistics park based on fuzzy clustering. In: 2009 second international conference on intelligent computation technology and automation, Changsha
13. Gao Y, Sun Y, Zheng W (2015) Genetic algorithm for optimizing the layout of logistics park based on shortest path and elitist strategy. J Inf Comput Sci 12(10):3765–3774
14. Zheng W, Gao Y (2015) Traffic impact analysis on interior logistics hub based on genetic algorithm of operation sequencing. J Inf Comput Sci 12(13):5197–5205
15. Hao H (2007) Working-area layout method study in logistics park based on genetic algorithm. Beijing Jiaotong University, Beijing

Route Choice Optimization for Urban Joint Distribution Based on the Two-Phase Algorithm

Qin Xiang

Abstract With the city's economic development, people have already put forward higher requirement for the city logistics, so joint distribution will attract increasing concern. In the daily operation, route choice is very important for the joint distribution, it will be related to the cost reduction and mitigation of congestion in urban transport. In order to solve the problem about the route choice in urban joint distribution network, the route choice model was established. Because normal algorithm would cost a lot of time, what' more, it also cannot get the optimal solution. Therefore, this paper provided the two-phase algorithm, which uses greedy algorithm to form the groups and apply ant colony algorithm for optimization. In order to verify the model and algorithm, through the case study, it shows that the unreasonable routes have already been adjusted, and the average line length has declined steadily. Compared with the result before, it decreases by 1.1%.

Keywords Joint distribution · Route choice · Two-phase algorithm · Greedy algorithm · Ant colony algorithm

1 Introduction

Joint distribution is also called as collaborative distribution, which was created in Japan. In the condition of increasingly congestion in urban traffic network, the joint distribution is a reasonable and advanced method which is developed in the process of practice. In 1992, by carrying out the joint distribution projects in Fujii area in Japan, the number of trucks decreased by 65%, the average transport distance for the truck everyday decreased by 28%, the total transport distance in this area decreased by 87%, the total parking frequency decreased by 72%, and total parking time was reduced by 17% [1]. In 1999, after the joint distribution system was carried out in Kassel in Germany, it can save the vehicle capacity by 80% [2].

Q. Xiang (✉)
Business School, Shanghai Normal University Tianhua College, Shanghai 201815, China
e-mail: 939606613@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_21

However, for the joint distribution, its service area is in the inner city, so its distribution distance is very short, and it also needs to turn over quickly. Although the total cost for joint distribution will consist of a lot of elements, the cost of trucks for delivery will have a large proportion, so in the process of operation, the route selection for the trucks would be very important, it is related to the total cost and quality of service. What is more, research on the vehicle route selection can not only improve the utilization and turnover rate of the vehicle so as to reduce the cost, but also it can ease urban traffic congestion and pressures of pollution.

2 Literature Review

For the vehicle route selection, Dantzig and Ramser initially proposed the VRP problem in 1959, they established the corresponding mathematical models and put forward the algorithm. In 1964, Clarke and Wright improved the method that was proposed by Dantzig and Ramser, then put forward the heuristic algorithm that was more effective [3]. In 1984, Golden et al. did some research on the multi-car vehicle selection [4]. In 1991, Desrochers et al. addressed the problem of simultaneously selecting the composition and routing of a fleet of vehicles in order to service efficiently customers with known demands from a central depot, then presented a new savings heuristic based on successive route fusion [5]. In 1999, Gendreau et al. described a tabu search heuristic for the Heterogeneous Fleet Vehicle Routing Problem (HVRP) [6]. In 2008, based on the consideration of the quality, Kun Chen proposed that there were two important targets for urban physical distribution route choice, which are minimizing the route travel time and maximizing the route travel time reliability, then established a route choice model [7].

From the literatures above, it shows that for the route choice model, there are lots of researches about that, and they have already had many achievements. But for the urban joint distribution, it is different from the ordinary vehicle routing problem (VRP), there are many terminal nodes in the region, and its service area are very large, so the problem about the route choice in urban joint distribution network belongs to large scale VRP problem. Although the algorithms, such as genetic algorithm, colony algorithm, neural network algorithm and so on, have outstanding performance in solving the small scale problem, for the large scale problem, they would cost lots of time. Therefore, in the process of solving this large scale problem, it should focus on the performance of the algorithm. In this paper, in order to solve the problem about the route choice in urban joint distribution, this paper established the model which considered the relationship between distribution center and terminal nodes, then according to this model, in order to speed up the operation, this paper designed the two-phase algorithm, which used greedy algorithm to form the groups and apply ant colony algorithm for optimization.

3 Route Choice Model

Before we establish the model, there are some assumptions as follow:

(1) In the model, the impacts that emergency brings will not be considered, such as weather, traffic accident, vehicle breakdown.
(2) The factors such as cargo shortage, inadequate vehicle capacity will not be considered.
(3) The space limitation of vehicle will not be considered.
(4) Because the distance from the distribution center to terminal nodes is not very long, maximum mileage of vehicle is not considered.
(5) The original and destination of vehicle are distribution center.
(6) Each terminal node can only be served by a truck once.

From the analysis above, it shows that the route choice problem in urban joint distribution belongs to the VRP problem that has single depot, multi-vehicle and load limitation. So in this paper, the author established the model to express the reality firstly.

In order to facilitate the description of the model, this paper uses the variables as follow:

c_{ij} The cost of vehicle from node i to node j
d_{ij} The distance from node i to node j
Q The capacity of vehicle
m The number of vehicles
n The number of nodes
d_i The weight of goods for node i
x_{ijk} If the truck k is from node i to node j, its value is 1, or its value is 0.
y_{ik} If the node i is served by truck k, its value is 1, or its value is 0.

In the daily operation, people always want to minimize the total vehicle distance or total cost. In this paper, the object is the minimum total cost, its formula is shown as follow:

$$\min \quad \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{m} c_{ij} x_{ijk} \tag{1}$$

From the assumption, each node can only be served once, so its formula is shown as follow:

$$\sum_{k=1}^{m} \sum_{i=0}^{n} x_{ijk} = 1 \tag{2}$$

In the daily operation, when the cargoes are loaded in the node, the truck will go to the next node soon, its formula is shown as follow:

$$\sum_{i=0}^{n} x_{ipk} - \sum_{j=0}^{n} x_{pjk} = 0 \tag{3}$$

For each node, it can only be served by a truck, so its formula is shown as follow:

$$\sum_{k=1}^{m} y_{ik} = 1 \tag{4}$$

For the truck, the amount of goods in a car cannot exceed the capacity of truck, so its formula is shown as follow:

$$\sum_{j=1}^{n} d_i y_{ik} \leq Q \tag{5}$$

For the delivery, each node can only be served by a truck that comes from other nodes, so its formula is shown as follow:

$$y_{jk} = \sum_{i=1}^{n} x_{ijk} \tag{6}$$

Therefore, by summing up of the analysis above, the model this paper establishes is shown as follow:

$$\begin{aligned} \min \quad & \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=1}^{m} c_{ij} x_{ijk} \\ st: \quad & \sum_{k=1}^{m} \sum_{i=0}^{n} x_{ijk} = 1 \\ & \sum_{i=0}^{n} x_{ipk} - \sum_{j=0}^{n} x_{pjk} = 0 \\ & \sum_{k=1}^{m} y_{ik} = 1 \\ & \sum_{j=1}^{n} d_i y_{ik} \leq Q \\ & y_{jk} = \sum_{i=1}^{n} x_{ijk} \end{aligned} \tag{7}$$

4 Algorithm Design for the Route Choice in Urban Joint Distribution

For the route choice model, it belongs to NP hard problem, using traditional method can only solve the small scale problem. And heuristic algorithm such as genetic algorithm and neural network algorithm can get the solution quickly for the a little large scale problem, but for the larger scale problem, it will have the problem of low efficiency and local optimum.

In the daily operation, each distribution center would provide the service to at least 200 terminal nodes, so the route choice belongs to a large scale problem. In this paper, considering the result performance for the model and calculation time, this paper proposes a two-phase algorithm, which uses greedy algorithm to form the groups and apply ant colony algorithm for optimization.

4.1 Greedy Algorithm

The definition of greedy algorithm is that when people solve the problem, they always select the solution that is considered to be the best currently. In this paper, the author selects greedy algorithm to form the group, so the greedy strategy taken in this paper is that when people load the goods, they consider the node which is the farthest from the distribution center firstly, then they would consider the node that is closed to the last node until the truck cannot load the goods, in this way, the group can be got for this truck. Its process is shown as follow:

(1) Initialize the network information, and input the number of the truck, its value is m.
(2) $i = 1$.
(3) If $i \leq m$, go to next step, or output the result.
(4) Select the farthest node from the distribution center, and load the goods, then update the truck capacity.
(5) Search the node that is closest to the last node and compare the weight of goods with the truck capacity, if the goods can be loaded, this node will be put into this group and update the network information, then go to step 6. If the goods cannot be loaded, go to step 7.
(6) Repeat step 5.
(7) $i = i + 1$, then go to step 3.

According to the process above, the groups can be got. During the process, the route can also be got initially. But because of the shortage of greedy algorithm, some routes are not reasonable, there will be a loop in the route which is shown in Fig. 1, so they need to be optimized. And in each group, it can be transferred to small scale travelling salesman problem (TSP) problem, using the heuristic algorithms can solve the problem easily and quickly. So the importance of greedy algorithm is transferring

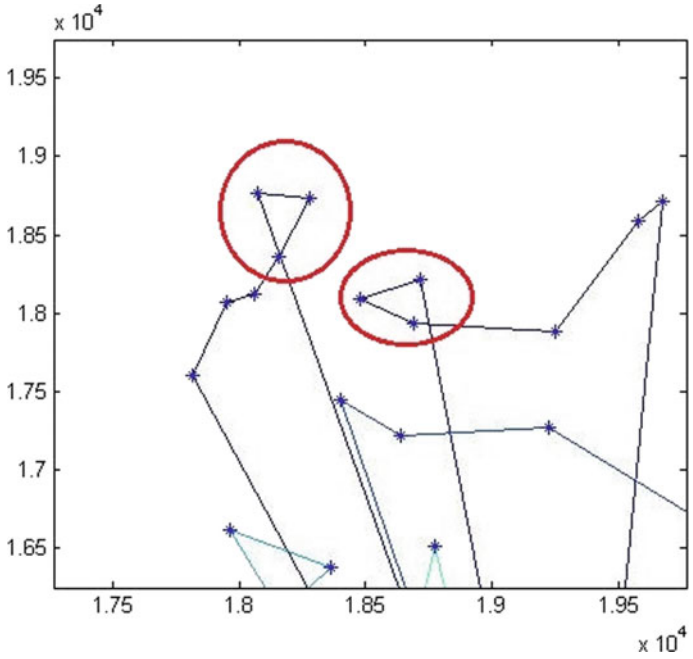


Fig. 1 The loop in a route using greedy algorithm

the route choice in urban joint distribution network that is large scale problem to some TSP problems that are small scale problem.

4.2 Ant Colony Algorithm

The ant colony algorithm is a probabilistic technique for solving computational problems which can be reduced to finding good paths through graphs. It was initially proposed by Marco Dorigo in 1992 in his Ph.D. thesis [8], the first algorithm was aiming to search for an optimal path in a graph, based on the behavior of ants seeking a path between their colony and a source of food. The original idea has since diversified to solve a wider class of numerical problems, and as a result, several problems have emerged, drawing on various aspects of the behavior of ants.

The TSP asks the following question: Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city? It is an NP-hard problem in combinatorial optimization, important in operations research and theoretical computer science. It is also a special case of the travelling purchaser problem and the Vehicle routing problem.

For the TSP, there are lots of algorithms to solve this problem. This paper choose ant colony algorithm. The process of ant colony algorithm is shown as follow:

(1) The shortest path is infinity.
(2) Initialization. All elements in the pheromone matrix released by all ants are zero, the city set that the ants have already visited is empty, all of cities can be visited. The starting positions for the ants are randomly selected, this node would be added in the set that is visited already and be deleted from the set that is allowed to visit.
(3) Select the next node for each ant. According to the Formula 8 as follow, each ant can select the next node, this node would be added in the set that is visited already and be deleted from the set that is allowed to visit. This step would be repeated until the set that is allowed to visit is empty. Then the pheromone matrix for each ant can be calculated according to Formula 9. Finally, the best route can be got, compared with the shortest path, if its distance is less than shortest path, this path will be the shortest path.

$$p_{ij}^{k} = \frac{(\tau_{ij})^{\alpha}(\eta_{ij})^{\beta}}{\sum_{j \in N}(\tau_{ij})^{\alpha}(\eta_{ij})^{\beta}} \tag{8}$$

p_{ij}^{k} The probability that ant k goes from city i to city j;
τ_{ij} Pheromone concentration from node i to node j;
η_{ij} A heuristic information. Its value is reciprocal of distance usually;
α Path visibility. Its value is always 1–2.
β: The weight of pheromone concentration. Its value is always 2–5.
N The set of cities that have never been visited and city i can reach.

$$\Delta\tau_{ij}^{k} = \frac{Q}{\sum L_k} \tag{9}$$

$\Delta\tau_{ij}^{k}$ Pheromone that is released by the ant k from node i to node j;
Q The total amount of pheromone released by an ant;
L_k The distance of ant k.

(4) Update the pheromone matrix according to the formula as follow:

$$\tau_{ij}(n+1) = \rho\tau_{ij}(n) + \sum_{k=1}^{m} \Delta\tau_{ij}^{k} \tag{10}$$

ρ Pheromone evaporation rate $0 \leq \rho \leq 1$

(5) Check the termination condition. If cycle number is maximum algebraic, go to step 6. If not, repeat the step 2, 3, 4.
(6) Output the result.

5 Case

In order to verify the feasibility of the model, this essay selected a distribution center as an example.

The distribution map of the terminal nodes is shown in Fig. 2. The density distribution is shown in Fig. 3.

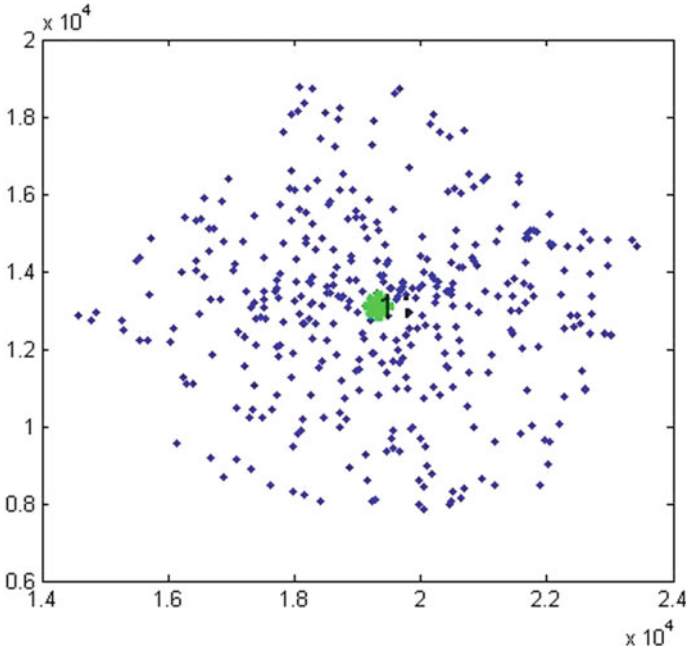


Fig. 2 Distribution map of the terminal nodes

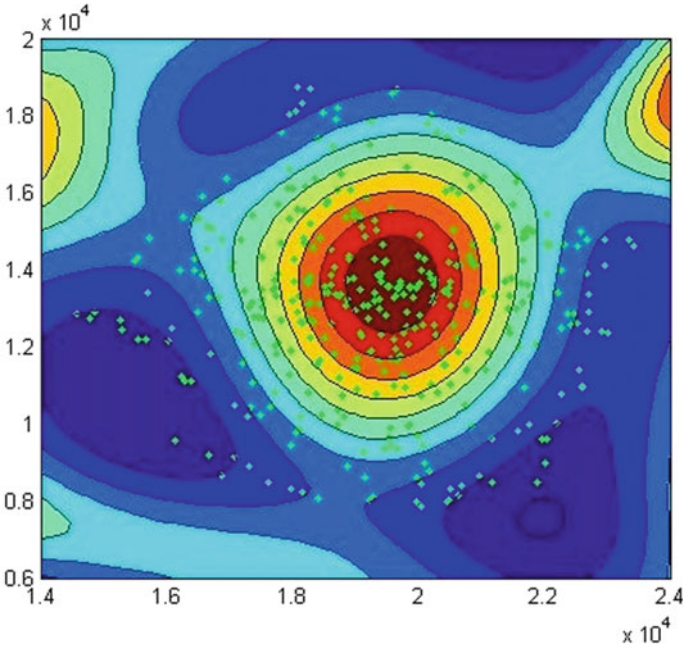


Fig. 3 Density distribution of the terminal nodes

According to the two-phase algorithm, firstly, this paper used greedy algorithm to transfer the route choice in urban joint distribution network that is large scale problem to some TSP problem that is small scale problem. Its result is shown in Fig. 4. From the figure, it shows that some paths are not reasonable, so they need to be optimized.

For the TSP, this paper used Ant colony algorithm for optimization. Its result is shown in Fig. 5. From the figure, it shows that the unreasonable paths have been adjusted. The iterative convergence chart is shown in Fig. 6. It can be seen from the figure that the average length of path shows decreasing trend, eventually it remains consistence. Compared with the result before, the length of best path decreases by 1.1%.

It needs to be clarified that in this paper, the distance between each node is Euclidean distance, but it doesn't affect the validity and stability of the two-stage algorithm. In the daily operation, the Euclidean distance can be replaced by the actual distance that can be got from GIS platform, the two-phase algorithm can still be suitable.

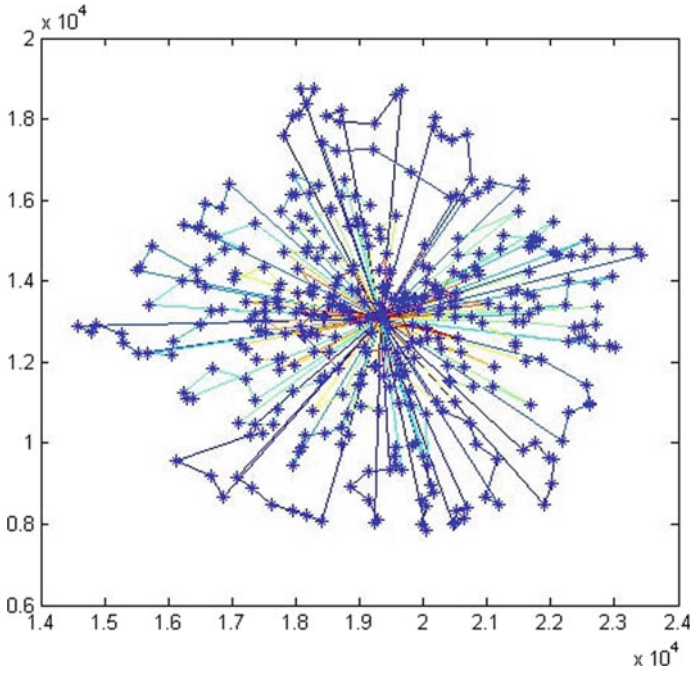


Fig. 4 Calculation result by greedy algorithm

6 Conclusion

For the route choice in urban joint distribution network, it belongs to large scale problem. In order to solve this problem, this paper established the route choice model. According to this model, as it is a NP-hard problem, and normal algorithm would cost a lot of time, what' more, it also cannot get the optimal solution. Therefore, this paper provided the two-phase algorithm, which uses greedy algorithm to form the groups and apply ant colony algorithm for optimization. In order to verify the model and algorithm, through the case study, it shows that the unreasonable routes have already been adjusted, and the average line length has declined steadily. Compared with the result before, it Decreases by 1.1%.

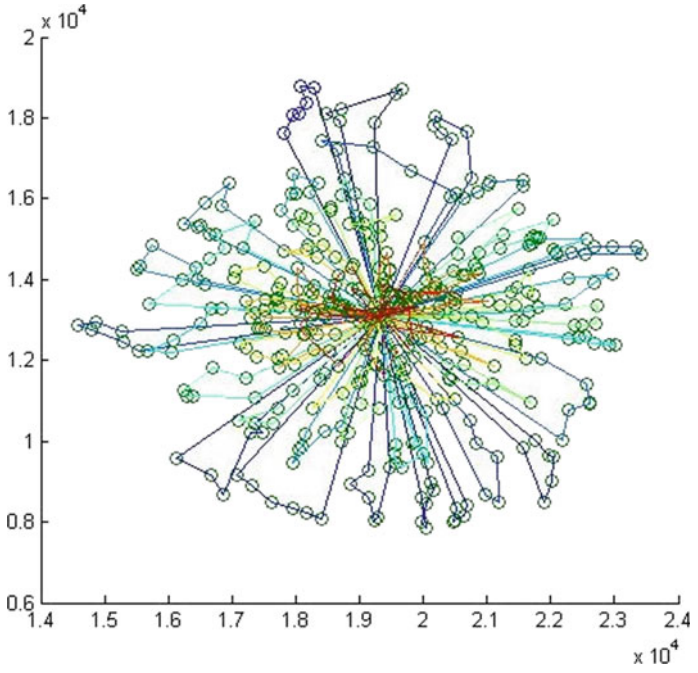


Fig. 5 Calculation result by ant colony algorithm

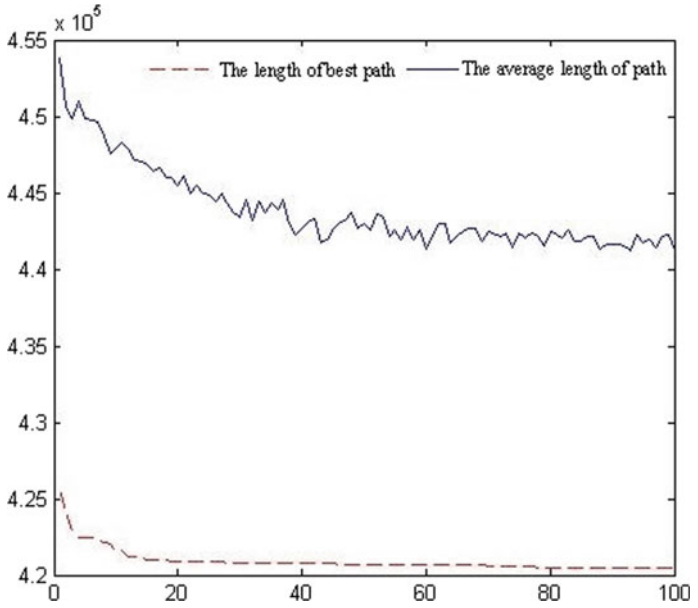


Fig. 6 Iterative convergence chart

References

1. Taniguchi (2002) Tenjin joint distribution system in Fukuoka. Best Practice Handbook, pp 7–12
2. Koehler U (1999) City logistics in Kassel. City Logistics 1:261–271
3. Clarke G (1964) Scheduling of vehicles from a central depot to a number of delivery points. Oper Res 12(4):568–581
4. Golden B, Assad A, Levy L et al (1984) The fleet size and mix vehicle routing problem. Comput Oper Res 11(84):49–66
5. Desrochers M, Verhoog TW (1991) A new heuristic for the fleet size and mix vehicle routing problem. Comput Oper Res 18(3):263–274
6. Gendreau M, Laporte G, Musaraganyi C et al (1999) A tabu search heuristic for the heterogeneous fleet vehicle routing problem. Comput Oper Res 26(12):1153–1173
7. Chen Kun, Geng Yanbin, Chai Dasheng (2008) Urban physical distribution route choice model based on travel cost. Logistics Technol 27(9):43–45
8. Dorigo M (1992) Optimization, learning and natural algorithms. Politecnico di Milano, Italy

An Integrated Energy-Efficient Scheduling and Train Control Model with Regenerative Braking for Metro System

Xinchen Ran, Shaokuan Chen and Lei Chen

Abstract Rising energy cost and environmental awareness make energy-efficient operation a key issue for metro management. The speed profile and timetable optimization are two significant ways to reduce total energy consumption for metro systems. This paper proposes an integrated speed profile and timetable optimization model to reduce the net energy consumption while incorporating with complex track conditions like undulate gradients, curves and tunnels. The net energy consumption is minimized by force coefficients and coast control for single train movement and accelerating and braking synchronization for multiple trains. An efficient hybrid particle swarm method based on the particle swarm optimization and genetic algorithm is designed to obtain a satisfactory solution. Finally, numerical case studies based on one metro line in Beijing are conducted to validate the energy-efficient performance of integrated model and the results show that the integrated model can achieve a better tradeoff between traction energy consumption and reused braking energy on comparison with individual speed profile and timetable optimization.

Keywords Metro system · Regenerative braking · Energy-efficient operation

1 Introduction

Due to the rising concerns on carbon emission and environmental problems, the energy efficiency of metro system has gradually become a key issue for operational management. In all types of energy-efficient measures, such as train trajectory and timetable optimization, energy storages system and reversible substations, train trajectory and timetable optimization are two effective and low-cost ways to save energy [1].

A considerable amount of literature has been published on train trajectory and timetable optimization. Train trajectory optimization is to find the energy-efficient

X. Ran · S. Chen (✉) · L. Chen
The MOE Key Laboratory of Transportation Complex Systems Theory and Technology, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
e-mail: shkchen@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_22

speed profile to drive the train through a specific section of track within a predetermined runtime [2]. The problem is usually formulated as an optimal control problem, aiming to find the regime sequence and corresponding switching points. The optimal control theory such as Pontryagin's maximum principle is used to prove that the optimal driving strategies should follow four regimes, which are maximum accelerating, cruising, coasting and maximum braking [3–7]. Based on these optimal regimes, sequence and switching points are optimized usually by numerical methods, such as dynamic programming [8, 9], genetic algorithm [10, 11] and ant colony algorithm [12].

However, train trajectory optimization usually considers single train movement while the utilization of regenerative energy is almost neglected. Therefore, the optimal speed profile for single train is not necessarily the optimal strategy for multiple trains [13]. For this reason, timetable optimization is widely studied to promote the utilization of regenerative energy by synchronizing the accelerating and braking trains. The problem can be divided into two levels: one is to optimize the timetable individually with the objectives of maximizing the overlapping time of accelerating and braking trains [14] or the total regenerative energy [15], which usually focus on the coordination of runtime, dwell time and headway, while the train trajectory is nearly predefined. Furthermore, another level is to optimize train trajectory and timetable jointly with the objective of minimizing net energy [16, 13, 17–19], i.e., the difference between traction energy consumption and reused braking energy. The heuristic algorithms especially genetic algorithm [16, 13, 17] has been widely used because of the high efficiency on computation time and solution quality. However, as the complexity of joint optimization is significantly improved, the integrated models usually make some simplification on train movement or track conditions, such as the maximum tractive or braking force of trains are assumed to be constants [16], the basic resistance of train is set as 0 [13, 17] or the additional resistances like curves and tunnel [18, 19] are neglected, which may have some influence on the practicability and accuracy of the results.

As an extension of above work, this paper proposes an integrated train speed profile and timetable optimization model incorporating with practical train movement and line conditions such as speed-dependent force control, undulate gradients, curves and tunnels. The model reduces the net energy by arranging the runtime, turnaround time, dwell time, force coefficients and switching points for multiple trains on the premise of service quality. The rest of this paper is organized as follows: Sect. 2 explains train movement and constructs the integrated optimal model. Section 3 proposes a hybrid particle swarm optimization (HPSO) algorithm to achieve a satisfactory solution. Section 4 demonstrated four case studies to explore the energy-efficient performance and Sect. 5 draws the conclusions.

2 Model Formulation

The problem is formulated as an integrated train trajectory and timetable optimization model considering multi-train movement and regenerative energy utilization for bi-directional metro lines. In this section, we first explain the basic formulas of train movement and energy consumption calculation for multiple trains. Then, the constraints of the integrated model are imposed. The assumptions are firstly set as follows:

Assumption 1: The speed profile on each section is assumed to follow the sequence of accelerating-coasting-braking and each regime adopts the same force coefficient.
Assumption 2: All trains are identical and considered as constant mass belts with uniform mass distribution.
Assumption 3: Each train complies with the same operation strategies at different cycles, i.e., the runtime and speed profile at each section, the dwell time and turnaround time at each station are all kept the same.

The metro line is illustrated as Fig. 1, where the number of stations is $2N$ and the number of sections is $2(N-1)$.

2.1 Train Movement

For a single-train journey between station n and station $n+1$, the runtime and energy consumption depend on the duration of accelerating, coasting and braking. As shown in Fig. 2, the speed profile for section n can be built by defining the tractive force coefficient μ_n^a, braking force coefficient μ_n^b and switching point x_n between accelerating and coasting.

According to Newton's laws of motion, the fundamental equation of train dynamic in section n between tractive force, braking force and resistance can be described as follows:

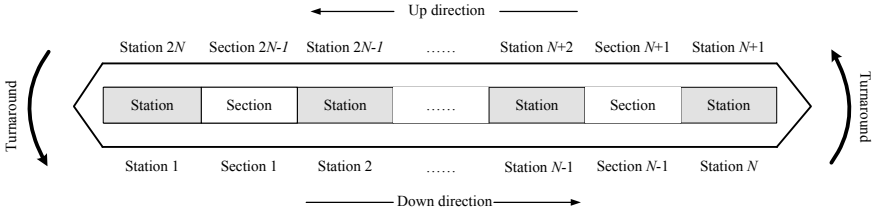


Fig. 1 An illustration on stations and sections of metro line

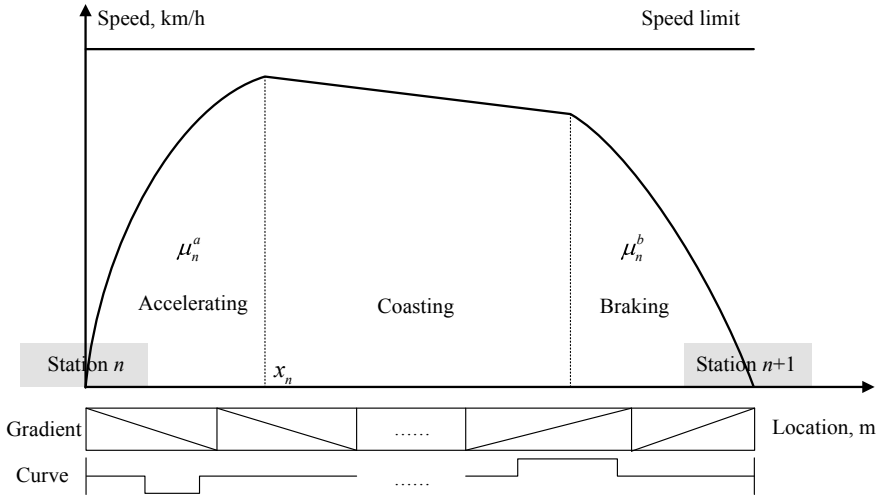


Fig. 2 An illustration on decision variables of speed profile

$$\rho M \frac{dv}{dt} = \begin{cases} \mu_n^a \cdot F_T(v) - R(v,x) & \mu_n^a \in (0,1] \\ -R(v,x) & \mu_n^a = \mu_n^b = 0 \\ \mu_n^b \cdot F_B(v) - R(v,x) & \mu_n^b \in [-1,0) \end{cases} \quad \forall n \in [1,N) \cup [N+1, 2N) \tag{1}$$

where M is the train mass; ρ is mass correction factor, usually set as 1.04 [20]; $F_T(v)$ and $F_B(v)$ are the maximum tractive and braking force applied to the train at speed v; $R(v,x)$ is the total resistance at location x with speed v, involving the basic vehicle resistance $R_v(v)$, gradient resistance $R_g(x)$, curve resistance $R_c(x)$ and tunnel resistance R_t.

$$R(v,x) = R_v(v) + R_g(x) + R_c(x) + R_t \tag{2}$$

$$R_v(v) = A + Bv + Cv^2 \tag{3}$$

$$R_g(x) = Mg\frac{i}{1000} \cdot \frac{L_s}{L_c} \tag{4}$$

$$R_c(x) = Mg\frac{600}{R} \cdot \frac{L_r}{L_c} \tag{5}$$

$$R_t = 0.00013MgL_t \tag{6}$$

The basis resistance $R_v(v)$ follows the Davis Eq. (3), where A, B and C are train-specific constants. On the basis of assumption 2, the gradient and curve resistance can be uniformly distributed to train by (4) and (5), where i is the gradient, positive

for uphill and negative for downhill; M is the mass of train; g is the gravity constant, set as 9.81 N/kg; L_s is the length of the train on the gradient; L_r is the length of train on the curve; L_c is the length of train. The tunnel resistance is determined by empirical Formula (6), where L_t is the length of tunnel.

2.2 Objective Function

The technology of using regenerative braking in metro systems could reduce the net energy consumption if the accelerating trains timely absorb the regenerative energy which is generated by the electric motor when trains are braking. Therefore, the objective of the integrated optimization model is to minimize the net energy consumption by synchronizing timetable and the accelerating and braking process for multiple trains. Firstly, we define the decision variables as timetable array DT and speed profile matrix $X_{(2N-2)\times 3}$, where D_n is the dwell time at station n; TR_1 and TR_{N+1} are the total turnaround times at terminal station 1 and $N+1$ respectively, including the dwell time at station 1, N, $N+1$ and $2N$.

$$DT = [TR_1, D_2, \ldots, D_{N-1}, TR_{N+1}, D_{N+2}, \ldots, D_{2N-1}] \tag{7}$$

$$X_{(2N-2)\times 3} = \begin{bmatrix} \mu_1^a, \mu_2^a, \ldots, \mu_{N-1}^a, \mu_{N+1}^a, \ldots, \mu_{2N-1}^a \\ \mu_1^b, \mu_2^b, \ldots, \mu_{N-1}^b, \mu_{N+1}^b, \ldots, \mu_{2N-1}^b \\ x_1, x_2, \ldots, x_{N-1}, x_{N+1}, \ldots, x_{2N-1} \end{bmatrix}^T \tag{8}$$

The total traction energy E_{tra} and regenerative energy E_{bra} can be calculated as the integral of mechanical power over runtime.

$$E_{tra} = \frac{1}{\varepsilon \cdot \rho_1} \sum_{n=1, n\neq N}^{2N-1} \sum_{k=1}^{K} \int_{t=t_{dn}^k}^{t_{a(n+1}^k} v(t) \cdot \mu_n^a \cdot F_T(v(t)) dt \tag{9}$$

$$E_{bra} = \frac{\rho_2}{\varepsilon} \sum_{n=1, n\neq N}^{2N-1} \sum_{k=1}^{K} \int_{t=t_{dn}^k}^{t_{a(n+1)}^k} v(t) \cdot |\mu_n^b| \cdot F_B(v(t)) \cdot \max\left(\frac{v(t) - v_l}{|v(t) - v_l|}, 0 \right) dt \tag{10}$$

where ε is the loss factor; ρ_1 is the efficiency from electrical energy to mechanical energy; ρ_2 is the efficiency from mechanical energy to electrical energy; t_{dn}^k is the departure time of train k from station n; $t_{a(n+1)}^k$ is the arrival time of train k at station $n+1$; K is the total train number. Note that the regenerative energy can be fed back to power grid only during the electrical braking process, the regenerative energy calculation should exclude the air braking process when the train speed is lower than critical speed v_l.

However, not all regenerative energy generated by braking trains can be reused. The utilization of regenerative energy should meet two conditions: one is that the braking trains and accelerating trains should be located in the same power supply area; the other is that the total energy consumption of accelerating trains is greater than regenerative energy, otherwise, the excess regenerative energy will be dissipated by on-board resistance. Consequently, the reused braking energy E_{rebra} should take the minimum value of total traction energy consumption and regenerative energy at instant t.

$$E_{rebra} = \sum_{q=1}^{Q} \int_{t=t_{dn}^{k}}^{t_{a(n+1)}^{k}} \{\min \left[\sum_{n=1,n\neq N}^{2N-1} \sum_{k=1}^{K} \frac{\lambda_{nq}}{\varepsilon \cdot \rho_1} v(t) \cdot \mu_n^a \cdot F_T(v(t)), \right.$$
$$\left. \sum_{n=1,n\neq N}^{2N-1} \sum_{k=1}^{K} \frac{\rho_2 \cdot \lambda_{nq}}{\varepsilon} v(t) \cdot \left|\mu_n^b\right| \cdot F_B(v(t)) \cdot \max\left(\frac{v(t)-v_l}{|v(t)-v_l|}, 0\right) \right] \Bigg\} dt \tag{11}$$

where λ_{nq} is a binary variable, if section n is located in substation q, λ_{nq}=1, otherwise, λ_{nq}= 0.

Finally, the net energy E_{net} could be calculated as the difference of total traction energy and total reused braking energy.

$$E_{net} = E_{tra} - E_{rebra} \tag{12}$$

2.3 Constraints

The model should satisfy the following constraints:

(1) Dwell time, runtime and turnaround time constraints

The dwell time, runtime and turnaround time should be integers and not exceed the upper and lower bounds.

$$D_n \in Z,\ \underline{D}_n \leq D_n \leq \bar{D}_n, \quad \forall n \in (1, N) \cup (N+1, 2N) \tag{13}$$

$$T_n \in Z,\ \underline{T}_n \leq T_n \leq \bar{T}_n, \quad \forall n \in [1, N) \cup [N+1, 2N) \tag{14}$$

$$TR_n \in Z,\ TR_n \geq \underline{TR}_n, \quad n = 1, N+1 \tag{15}$$

(2) Headway and cycle time constraints

The headway h_0, train number K and cycle time T_{cycle} should be integers and satisfy (16). In addition, the headway and cycle time remain unchanged to reduce the impact on rescheduling and service quality.

$$T_{cycle}, K, h_0 \in Z, \quad T_{cycle}/K = h_0 \tag{16}$$

$$\sum_{n=1,n\neq N}^{2N-1} T_n + \sum_{n=2,n\neq N,N+1}^{2N-1} D_n + TR_1 + TR_{N+1} = T_{cycle} \tag{17}$$

(3) Comfort constraint

In order to ensure the comfort of passengers, the acceleration of train should not exceed permissible acceleration $\bar{\alpha}$ and deceleration $\bar{\beta}$, set as $\bar{\alpha} = 1\,\text{m/s}^2, \overline{\beta} = -0.83\,\text{m/s}^2$ [21].

$$\bar{\beta} \leq dv(t)/dt \leq \bar{\alpha} \tag{18}$$

(4) Speed constraint

The speed of train at any moment should not exceed the speed limit and the speed at station mileage S_n should be 0.

$$0 \leq v(t) \leq \bar{v} \tag{19}$$

$$v(S_n)=0, \quad \forall n \in [1, N) \cup [N+1, 2N] \tag{20}$$

(5) Tractive and braking force constraint

The tractive and braking force should not exceed the maximum force. For convenience, the force coefficients are dispersed to 10 levels.

$$10\mu_n^a,\ 10\mu_n^b \in Z,\ 0<\mu_n^a \leq 1, -1 \leq \mu_n^b < 0, \quad \forall n \in [1, N) \cup [N+1, 2N) \tag{21}$$

(6) Switching point constraint

The switching point from accelerating to coasting x_n should be an integer and located between station mileage S_n and S_{n+1}.

$$x_n \in Z,\ S_n < x_n < S_{n+1}, \quad \forall n \in [1, N) \cup [N+1, 2N) \tag{22}$$

Based on above analysis, the integrated energy-efficient operation model can be constructed as follows:

$$\begin{aligned} &\min E_{net} = E_{tra} - E_{rebra} \\ &s.t. \quad (1)-(11), \quad (13)-(22) \end{aligned} \tag{23}$$

3 Solution Algorithm

This section designs a hybrid particle swarm optimization (HPSO) algorithm to solve the integrated energy-efficient operation model and then the key steps and flowchart of the HPSO algorithm are introduced.

The HPSO algorithm is the combination of particle swam optimization algorithm and genetic algorithm. The crossover and mutation are performed after particle updating to prevent the algorithm falling into local optimum. The first step for HPSO algorithm is to construct a one to one mapping between solution space and particle structure. Figure 3 shows the real number coding structure of a particle, which is composed by a decision matrix with 4 rows and $2(N-1)$ columns. Each row represents turnaround time and dwell time, tractive force coefficients, braking force coefficients and switching points successively.

The key steps for HPSO algorithm are explained as follows:

Step 1: Initialization. Each particle P is initialized with uniform random value between the lower and upper boundaries satisfying (13), (15), (15), (21) and (22). Moreover, the speed V of each particle is randomly initialized within the range of $[-1, 1]$.

Step 2: Evaluation function calculation. The evaluation function is the sum of objective function (12) and penalty function, consisting of cycle time, acceleration and speed limit constraints. The objective function can be calculated by train simulation module, as shown in Fig. 4.

$$Eva = E_{net} + p_1(T'_{cycle} - T_{cycle}) + p_2\sigma_1 + p_3\sigma_2 \tag{24}$$

where T'_{cycle} is the actual cycle time of each particle; σ_1 and σ_2 are binary variables, if the particle satisfies (18), $\sigma_1 = 0$, otherwise, $\sigma_1 = 1$; if the particle satisfies (19), $\sigma_2 = 0$, otherwise, $\sigma_2 = 1$; p_1,p_2 and p_3 are the penalty factors of cycle time, comfort and speed limit, usually set as large constants.

	Down direction				Up direction			
Turnaround time dwell time	TR_1	D_2	$\cdots$	D_{N-1}	TR_{N+1}	D_{N+2}	$\cdots$	$D_{2N\text{-}1}$
Tractive force coefficients	μ_1^a	μ_2^a	$\cdots$	μ_{N-1}^a	μ_{N+1}^a	μ_{N+2}^a	$\cdots$	$\mu_{2N\text{-}1}^a$
Braking force coefficients	μ_1^b	μ_2^b	$\cdots$	μ_{N-1}^b	μ_{N+1}^b	μ_{N+2}^b	$\cdots$	$\mu_{2N\text{-}1}^b$
Switching points	x_1	x_2	$\cdots$	x_{N-1}	x_{N+1}	x_{N+2}	$\cdots$	x_{2N-1}

Fig. 3 Particle structure

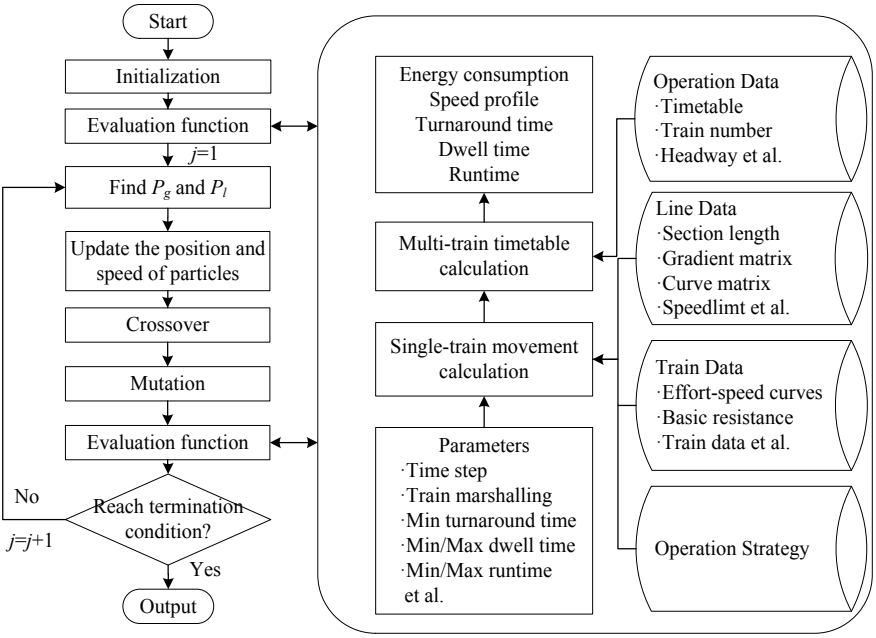


Fig. 4 The flowchart of HPSO algorithm

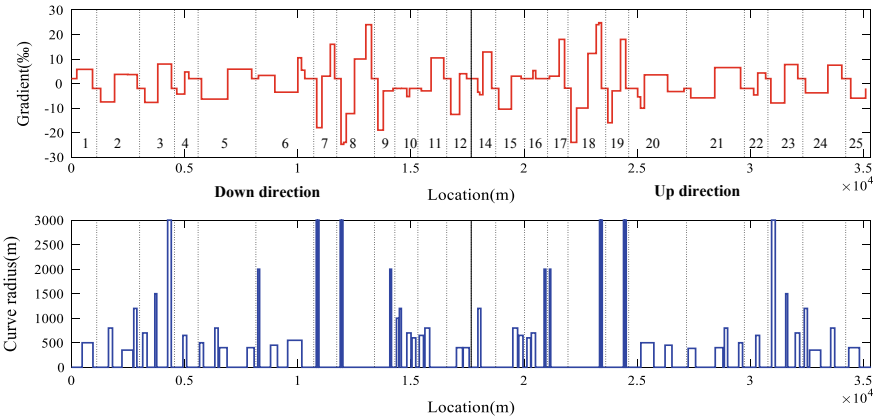


Fig. 5 Distribution of gradients and curves

Step 3: Find two best particles. The first particle is the one with best evaluation value which has achieved so far in all past iterations, defined as P_g. The second particle is the one with best evaluation value in current iteration, defined as P_l.
Step 4: Update the position and speed of particles. According to the following equations, the particles P update themselves by P_g and P_l, where γ is the inertia weight;

c_1, c_2 are two learning factors; r_1, r_2 are random numbers within the range of [0,1].

$$V(j+1) = \gamma \cdot V(j) + c_1 \cdot r_1 \cdot (P_g(j) - P(j)) + c_2 \cdot r_2 \cdot (P_l(j) - P(j)) \tag{25}$$

$$P(j+1) = P(j) + V(j) \tag{26}$$

Step 5: Crossover. Cross the particles and P_g with the probability of P_c by using two-point crossover strategy. If new particles do not satisfy constraints in (23), a new crossover operation will be performed again.
Step 6: Mutation. Select one particle with the probability P_m and change the value at a random position. Validity examination also has to be done as **Step 5**.
Step 7: Evaluation function calculation. Update the evaluation values of new particles as **Step 2**.
Step 8: Check termination criterion. If the algorithm has reached predetermined maximum number of iteration, output dwell time, runtime, turnaround time and driving strategies of P_g, otherwise, go back to **Step 3**. The flow chart of the HPSO algorithm is shown in Fig. 4.

4 Experimental Study

In order to evaluate the energy-efficient performance of integrated optimization model, this section conducts numerical experiments based on one metro line in Beijing and analyzes the effect of different optimal strategies.

4.1 Set up

The metro line covers a length of 17.7 km and contains 13 stations. Each train set is composed of 6 cars with a mass of 282.08t and a length of 118.8 m. The tractive and braking effort-speed curves of each motor are piecewise functions, as shown in Fig. 6, where the critical speed of regenerative braking and air braking is 5 km/h. The Davis equation coefficients A, B, C are set as 2.378, 0.028, 0.657×10^{-3}. The station mileage and distribution of substations are shown in Table 1 and the distributions of gradients and curves are shown in Fig. 5.

In this paper, the multi-train movement during one peak hour for both directions is studied. The headway is set as 180 s, the cycle time is 3960 s, the train number is 22 and the minimum turnaround time at each terminal station is set as 150 s. In the integrated model, parameters are set as follows: loss factor $\varepsilon = 0.1$, the efficiency from electrical energy to mechanical energy $\rho_1 = 0.9$, the efficiency from mechanical energy to electrical energy is $\rho_2 = 0.6$ [16]. For HPSO algorithm, the initial particle

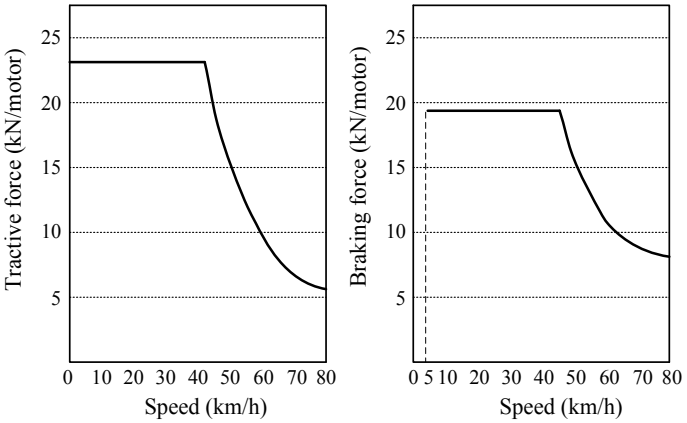


Fig. 6 Tractive and braking effort-speed curves of each motor

number is 50, the learning factors c_1, c_2 are both set as 1.45, inertia weight γ is 0.8, the crossover probability P_c is 0.8, mutation probability P_m is 0.05. In multi-train simulation, the time step is set as 0.1 s.

4.2 Result and Analysis

In order to compare the optimal strategies and evaluate the effect of force coefficients, we design four cases on the basis of original timetable: Case 1 only optimizes the speed profile in each section; Case 2 only optimizes the timetable using the same speed profile with Case 1; Case 3 optimizes the speed profile and timetable jointly and Case 4 changes the force coefficients but the timetable is kept the same with Case 3.

The dwell time, turnaround time and runtime of each case are listed in Table 2. Compared with the individual speed profile optimization (Case 1) and timetable optimization (Case 2), the compressive optimization strategies (Case 3 and Case 4) increase the runtime of most sections and compresses the dwell time at most stations. Figure 7 shows the speed profiles of each case, where both individual optimization (Case 1 and Case 2) and compressive optimization (Case 3) follow the sequence of maximum accelerating-coasting-maximum braking. Moreover, the duration of accelerating and braking process can be extended in two different ways: one is to compress the runtime of section (compare Case 1/2 to Case 3 in Sect. 4–14, 17, 19–22, 25), the other is to drop the force coefficients of accelerating and braking (compare each section of Case 3 and Case 4).

The comparisons on energy performance of each case are listed in Table 3, where the traction energy, regenerative energy, reused braking energy and net energy are calculated by (9), (10), (11) and (12) respectively. Compared with the individual

Table 1 Station mileage and substation distribution

Down direction	1	2	3	4	5	6	7	8	9	10	11	12	13
Up direction	26	25	24	23	22	21	20	19	18	17	16	15	14
Station mileage	0	1114	3009	4552	5594	8147	10705	11721	13389	14289	15307	16582	17665
Substation	I	I	I	I/II	II	II	II/III	III	III	III/IV	IV	IV	IV

Table 2 Timetable of each case

Down direction	Number	1	2	3	4	5	6	7	8	9	10	11	12
Case 1	Dwell time	230	45	35	30	40	30	30	30	30	60	40	40
Case 1/2	Run time	92	142	113	85	177	175	86	126	79	87	98	89
Case 2	Dwell time	250	40	30	25	35	35	29	25	35	65	35	35
Case 3/4	Dwell time	246	40	30	25	35	25	25	35	25	55	35	35
Case 3/4	Run time	97	137	109	90	182	180	91	129	84	89	99	94
Up direction	Number	14	15	16	17	18	19	20	21	22	23	24	25
Case 1	Dwell time	230	40	40	60	30	30	30	30	40	30	30	45
Case 1/2	Run time	91	95	85	80	118	83	177	177	85	113	140	92
Case 2	Dwell time	252	39	35	55	35	25	35	35	35	25	25	40
Case 3/4	Dwell time	245	35	35	55	35	35	25	25	35	25	25	40
Case 3/4	Run time	95	93	83	83	117	88	182	182	89	108	136	97

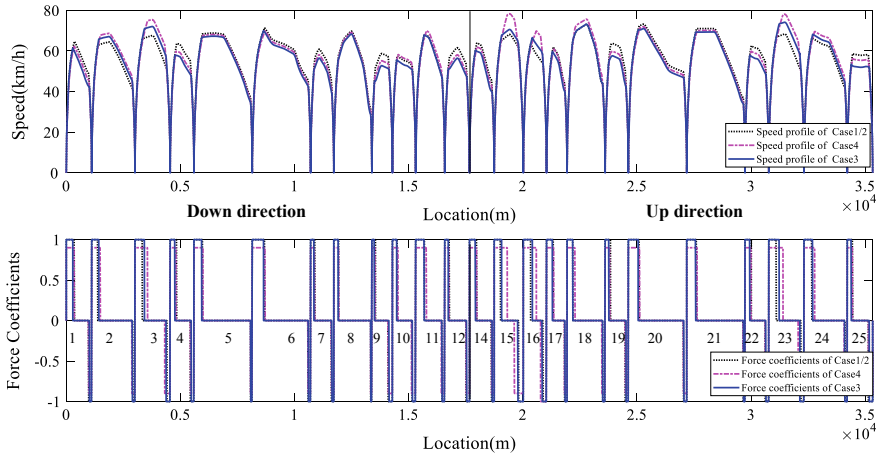


Fig. 7 Operation strategies of each case

Table 3 Comparison on energy performance of each case

Index	Cycle time	Dwell time	Run time	Traction energy	Regenerative energy	Reused energy	Net energy
Case 1	3960	1275	2685	8938.14	2139.54	1060.88	7877.26
Case 2	3960	1275	2685	8938.14	2139.54	1613.07	7325.07
Case 3	3960	1226	2734	8538.45	1968.86	1749.78	6788.67
Case 4	3960	1226	2734	9234.03	2273.96	2024.03	7210.01
Case 2 versus Case 1	0	0	0	0.00%	0.00%	52.05%	−7.01%
Case 3 versus Case 1	0	−49	49	−4.47%	−7.98%	64.94%	−13.82%
Case 4 versus Case 1	0	−49	49	3.31%	6.28%	90.79%	−8.47%

optimization strategies Case 1, Case 2 and the strategy with lower force coefficients (Case 4), the comprehensive model Case 3 achieves the best energy conservation.

In addition, Case 3 lengthens the runtime from 2685 to 2730 s in comparison with Case 1 and Case 2, which can significantly reduce the traction energy consumption. Although an influence on total regenerative energy is happened arising from the reduction on braking duration, the reused braking energy is totally increase by timetable arrangement. Conversely, Case 4 increases the reused braking energy utmostly by extending the duration of accelerating and braking, but it consumes more traction energy, resulting in more net energy consumption than Case 3. Therefore,

the integrated model can get a better tradeoff between traction energy consumption and reused braking energy.

5 Conclusion

In this paper, an integrated speed profile and timetable optimization is proposed to minimize the net energy consumption for metro system and the effectiveness of the method is carried out by numerical case studies based on one metro line in Beijing. The results of different cases show that the integrated optimization model can make a better tradeoff on traction energy consumption and reused braking energy on comparison with individual speed profile or timetable optimization. Moreover, the optimal train movement strategy of integrated model should adopt individual optimal speed profile, which follows the sequence of maximum accelerating, coasting and maximum braking. Although lower force coefficients could contribute to higher utilization of regenerative energy, it consumes more traction energy. However, note that the speed profile adopt in this paper only contains three regimes, more complicated regime sequence adapting to different track conditions like long-distance section and variable speed limits will be studied further to investigate the impact on traction energy consumption and utilization of regenerative energy.

Acknowledgements The authors are grateful to the National Natural Science Foundation of China (71571015, 71621001) for their financial supports.

References

1. Douglas H, Roberts C, Hillmansen S et al (2015) An assessment of available measures to reduce traction energy use in railway networks. Energy Convers Manag 106(12):1149–1165
2. Scheepmaker GM, Goverde RMP, Kroon LG (2017) Review of energy-efficient train control and timetabling. Eur J Oper Res 257(2):355–376
3. Howlett P (1990) An optimal strategy for the control of a train. Anziam J 31(4):454–471
4. Khmelnitsky E (2000) On an optimal control problem of train operation. Autom Control IEEE Trans 45(7):1257–1266
5. Liu R, Golovitcher IM (2003) Energy-efficient operation of rail vehicles. Transp Res Part A 37(10):917–932
6. Albrecht A, Howlett P, Pudney P et al (2016) The key principles of optimal train control—Part 1: formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points. Transp Res Part B Methodol 94(2):482–508
7. Albrecht A, Howlett P, Pudney P et al (2016) The key principles of optimal train control—part 2: existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques. Transp Res Part B Methodol 94(2):509–538
8. Lu S, Hillmansen S, Ho TK et al (2013) Single-Train trajectory optimization. IEEE Trans Intell Transp Syst 14(2):743–750
9. Zhao N, Roberts C, Hillmansen S et al (2015) A multiple train trajectory optimization to minimize energy consumption and delay. IEEE Trans Intell Transp Syst 16(5):2363–2372

10. Chang CS, Sim SS (1997) Optimising train movements through coast control using genetic algorithms. Electric Power Appl IEE Proc 144(1):65–73
11. Wong KK, Ho TK (2004) Dynamic coast control of train movement with genetic algorithm. Int J Syst Sci 35(13–14):835–846
12. Ke BR, Chen MC, Lin CL (2009) Block-layout design using MAX–MIN ant system for saving energy on mass rapid transit systems. IEEE Trans Intell Transp Syst 10(2):226–235
13. Yang X, Li X, Ning B et al (2015) A survey on energy-efficient train operation for urban rail transit. IEEE Trans Intell Transp Syst 17(1):2–13
14. Yang X, Li X, Gao Z et al (2013) A cooperative scheduling model for timetable optimization in subway systems. IEEE Trans Intell Transp Syst 14(1):438–447
15. Pena-Alcaraz M, Fernandez A, Cucala AP et al (2011) Optimal underground timetable design based on power flow for maximizing the use of regenerative-braking energy. Proc Inst Mech Eng Part F J Rail & Rapid Transit 226(4):397–408
16. Li X, Hong KL (2014) An energy-efficient scheduling and speed control approach for metro rail operations. Transp Res Part B Methodol 64(4):73–89
17. Yang X, Chen A, Li X et al (2015) An energy-efficient scheduling approach to improve the utilization of regenerative energy for metro systems. Transp Res Part C Emerg Technol 57:13–29
18. Ye H, Liu R (2016) A multiphase optimal control method for multi-train control and scheduling on railway lines. Transp Res Part B Methodol 93:377–393
19. Zhao N, Roberts C, Hillmansen S et al (2017) An integrated metro operation optimization to minimize energy consumption ☆. Transp Res Part C Emerg Technol 75:168–182
20. Chevrier R, Pellegrini P, Rodriguez J (2013) Energy saving in railway timetabling: a bi-objective evolutionary approach for computing alternative running times. Transp Res Part C Emerg Technol 37(3):20–41
21. Ding Y, Liu H, Bai Y et al (2011) A two-level optimization model and algorithm for energy-efficient urban train operation. J Transp Syst Eng Inf Technol 11(1):96–101

A Review of the Research on the Entrance Control Method of Urban Expressway

Yan Xing, Jin-ling Wang, Wei-dong Liu, Xing-quan Guan and Yang Liu

Abstract Expressway control is an important measure to solve the congestion of urban expressway nodes and ensures the smooth and efficient operation of expressway lines through the management and regulation of traffic volume. In order to further study the control of expressway, it is necessary to classify and summarize the existing entrance control model. According to the control method, control range, control effect and control complexity, the application model of urban expressway entrance control is classified and summarized according to static, dynamic single-point, multi-ramp coordinated control, internal integrated control and integrated control of expressway and auxiliary road intersection. Then, bases on existing research theories and practices, clears model application optimization technology and modeling ideas, analyzes the advantages and disadvantages of various basic models and improves models. Finally, the urban expressway control is prospected.

Keywords Highway transportation · Ramp metering · Coordinated control · Urban expressway · Traffic flow characteristics

1 Introduction

The expressway is an essential channel connecting the central area of the city and the outskirts of the city. It has attracted people's attention due to its high transportability, strong connectivity, fast speed, and so on. And it is widely used in urban traffic planning. However, with increasing traffic demand, the advantages of expressways are limited. Entrance control is the main point of the expressway system control, which determines the operating status of the traffic flow on the expressway. It is an important aspect of the expressway planning. Scientific and reasonable control can

J. Wang · W. Liu (✉) · X. Guan · Y. Liu
School of Transportation Engineering,
Shenyang Jianzhu University, Shenyang 110168, Liaoning, China
e-mail: 44667276@qq.com

Y. Xing
School of Automotive Engineering, Jilin University, Changchun 130025, Jilin, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_23

reduce the influence of the entrance traffic flow on the main road traffic flow, improve the efficiency of expressway transportation, ensure a balanced distribution of main road traffic, achieve the optimization of system efficiency, and satisfy people's traffic demand for expressways.

Expressways and highways are long-distance, large-capacity, rapid transit, and the operating environments are relatively closed. At present, most of the studies on expressway are based on the researches of the highway. However, in this more complex operating environment, due to the frequent conversion of urban expressway traffic flow, and the control model of the highway considers less control parameters, the effect is not ideal in the application to the expressway. Therefore, the model needs to be improved. There are many connecting points between the expressway and the ground, which are connected with different grades of roads between cities. When the road is connected, the speed of the road is limited by the state of the road and traffic flow. And the traffic flow is large in a short distance, so the traffic contradiction between the entrance ramp and the main road interlacing is prominent.

Although the research on the expressway entrance is continuously improving, due to the unpredictable real-time nature of traffic flow, various models cannot fully consider their influencing factors and are often limited in practical applications. This paper systematically summarized the expressway entrance control model, analyzed and integrated the advantages and shortcomings of the model, proposed research ideas for the future expressway entrance control. Finally, according to the current application of dynamic model, the prospect of expressway entrance control is forecasted.

2 Entrance Control Method

The study of expressway entrance control began in the 1960s. The earliest control application on the ramp was in Chicago, United States. The management of traffic flow at the entrance restricted the inflow of vehicles and alleviates the traffic congestion at the entrance interchange to some extent. As the theory continues to enhance, more and more models have been established for expressway entrance control. In this paper, the differences among the static control model, dynamic single- point model and dynamic coordination model are classified and analyzed.

2.1 Static Models

As the problem of traffic congestion at the expressway entrance became increasingly prominent, the timing control method was proposed by Watleworth in 1965 [1]. The method limits the flow of traffic on the ramp for problems such as high inlet flow at peak times and large traffic impacts on the main road. And by utilising the existing traffic data, we can determine the maximum traffic flow that the main road

can accommodate and the ramp release. It is based on the main road traffic flow states to determine the interval between the release of time. The use of this timing release method is designed to divide the continuous flow at the entrance of the ramp into intermittent flow into the main road traffic.

When timing control does not consider the actual traffic demand of the entrance, Papageorgiou proposed a new control method in 1980 [2]. The method takes into account the degree of change in traffic flow over time and gives the NK constraints of the static model. It is based on a linear optimization optimal solution and capacity constraints to establish a model. Although this model takes into account the changes in traffic flow, it can adjust the traffic flow in the peak state to a stable state. However, it is limited to the timing control and cannot handle real-time traffic flow flexibly.

2.2 On-Ramp Single Point Control Dynamic Model

With the development of technology and the ever-increasing demand for actual traffic, dynamic control methods have been proposed for the problem of poor adaptability and limitation of static control. The first proposed dynamic model is a single-point of control. It is an on ramp single-point control, which is the signal control of an independent ramp entrance. According to the real-time traffic flow status at the current entrance, the ramp regulation rate is determined and the entrance control is realized. In the single-point control, the classic is the demand-capacity difference algorithm, ALINEA algorithm, and single-channel control intelligence algorithm.

The demand-capacity control method is an open-loop control strategy and has weak ability to handle external disturbances. By introducing an on-ramp dynamic regulation system and a control algorithm with poor traffic demand and traffic capacity, Lan et al. [3] proposed the issue of selecting the on-ramp dynamic regulation rate and studied the dynamic adjustment of the on-ramp metering channel. Hu [4] comprehensively analyzed the advantages and disadvantages of the demand-capacity balance control algorithm and the ALINEA algorithm, then established a new model. That improved the import capacity of the on-ramp and the service level of the main road on the basis of maximizing the use of expressways. Liu et al. [5] used the speed as the control parameter based on the capacity control method. This method reduces the average delays of ramps and road networks and increases the average speed downstream of ramps to the fastest.

ALINEA is a single-point control model with wide application and classic closed-loop feedback control [6]. At present, various dynamic models are modifications based on the ALINEA model or revisions to their improved models. The ALINEA algorithm utilizes the on-ramp adjustment rate of the previous time period $r(k-1)$, the occupancy rate of the downstream main road $O_{out}(k)$, and the occupancy rate of the downstream ideal state O_d, and smoothes the occupancy difference between the two periods (Eq. 1).

The basic model is:

$$r(k) = r(k-1) + K_R[O_d - O_{out}(k)] \tag{1}$$

The constraint is:

$$\max[d(k-1) - (L_{\max} - \frac{l(k)}{T}, r_{\min}) \leq r(k) \leq \min[d(k-1) + \frac{l(k)}{T}, r_{\max}] \tag{2}$$

where K_R is the adjustment parameter; $l(k)$ is the inlet queue length.

Liu [7] and Chi [8] proposed a fuzzy self-tuning control method for the entrance of ALINEA model to suppress the disturbance of the external environment to the traffic flow. Fang [9] used the vehicle insertable gap as a basis for the calibration of the value of the critical occupancy rate, and established the ramp hierarchical control based on the relevant queuing experience, which had a good effect on the main road traffic.

Some scholars have applied intelligent algorithms to the control model in single lane control in order to effectively control the complex, variable, and nonlinear traffic flow of the expressway. Yan [10] studied the fuzzy control of expressway entrances, and confirmed the effectiveness and real-time nature of fuzzy logic control. Zhang et al. [11] applied dynamic fuzzy neural network to the entrance control of ramps to improve the stability of the neural network algorithm.

2.3 Multi-ramp Coordination Dynamic Models

On-ramp coordinating control is the real-time monitoring of multiple ramps, taking into consideration the main road and ramp capacity of each segment, traffic flow status, and so on. It is aimed at the optimality of the overall planning of the expressway through the control of the adjustment rate of the entrance ramps, and achieve the overall coordination control of multi ramps. In the multi-channel coordinated control, the METANET model and the extended model of ALINEA are more common.

The METANET model is a macroscopic dynamic traffic flow model. The discretization of time and road sections improves the accuracy and applicability of the model. It has been improved and applied in the study of coordination control in the expressway. Hu [12] based on the METANET model, according to the mutual constraints between variables to give numerical solutions, and select multi-channel nodes in the MPC framework, taking into account the environmental benefits and the benefits of traffic to establish a multi-ramp dynamic coordination model. Cao [13] proposed a queuing length estimation method based on improved input-output model and shock wave model according to the changes in the number and density of inflow and outflow vehicles. Then, the Vissim software was used to verify that

the shock wave model-based estimation method is more suitable for multi-channel coordinated entry control strategies.

METALINE [14] is a coordinated feedback-based ramp control strategy proposed for the extension of the ALINEA model with high stability and sensitivity. It introduces the vector combined with smoothness to solve the regulation rate of each ramp entrance, and derives the multivariable control formula: Physical control law (LQ) (Eq. 3) and Linear quadratic integral (LQI) control law (Eq. 4). Gao [15] optimized the multi-lane entry ramp and proposed an on-ramp density dynamic equation. Considering that the proportion of traffic flow in the inner and outer lanes of the multiple ramps is quite different, the ALINEA model was improved and the effectiveness of the model in application was improved.

$$\underline{r}(k) = \underline{r}^{d} - K_{LQ}[\underline{\rho}(k) - \underline{\rho}^{d}] \tag{3}$$

$$\underline{r}(k) = \underline{r}(k-1) - K^{1}_{LQI}[\underline{\rho}(k) - \underline{\rho}(k-1)] - K^{2}_{LQI}[\underline{\hat{\rho}}(k) - \underline{\rho}_{d}] \tag{4}$$

where $\underline{r}(k)$ is controllable slope volume vector; $\underline{\rho}(k)$ is density vector; $\underline{\hat{\rho}}(k)$ is the vector density of certain selected bottlenecks; K is the time index.

In addition to the above-mentioned multi-ramp coordinated control methods, there is a fuzzy neural network control method based on fuzzy calculation and neural network theory [16], and a switching control method based on the hybrid automaton model [17], based on multi-ramp control Multi-agent reinforcement learning control system [18] and the control model based on PID neural network multivariable control system [19].

2.4 Integrated Control Model in the Expressway

The integrating control of the expressway system refers to the expansion of the local small-scale optimization into the coordinated optimization based on the correlation between the various components within the expressway system and its own characteristics, so that the overall control effect of the system is optimized. Among them, the TBC model and the MetaNet model are widely used, and many scholars have studied and improved it.

The TBC model combines the control model of the multi-machine parallel system (MPP system) with the ALINEA model to achieve the transfer and balance of traffic demand. It improved the local congestion situation, but the application effect is not satisfactory. Many scholars improved the model. Zhu et al. [20] controlled the vehicle differential value based on the entrance control method of task balance control (TBC) for the uneven flow distribution on the main and auxiliary roads. To some extent, it satisfies the traffic demand of the vehicle, but this method only considers the vehicle adjustment in the local area, and the overall regulation effect on the expressway main road traffic is not obvious. Zhu et al. [21] improved the model (TBC) and proposed an

entrance balanced control model (TBM) based on the main idea of the utility balance of main and auxiliary road resources, introducing the difference parameter 'd' and interference factor 'f' results in an asymmetry-improved model and an improved time-space resource equalization model.

Due to its high accuracy and applicability, the METANET model is gradually applied to in-flight aggregate control. Lu [22] revised the MetaNet model in 2011 and proposed that ramp-variable speed limit integrated control can effectively alleviate traffic bottlenecks. Sun [23], Zhu et al. [24] used the MPC prediction model and ant colony algorithm respectively to optimize the MetaNet model and improve the vehicle's operating efficiency. Liang [25] comprehensively considered the MetaNet model and variable-constrained coordinated control, studied and analyzed the interaction between the on-ramp control and the variable speed limit control, maximized traffic efficiency and environmental benefits while ensuring maximum road network efficiency.

Some scholars have proposed different control methods for the integrated control of expressways. Yang [26] proposed the concept of coordinated control of on-ramp and main roads, and used models to analyze the influence of different factors on road traffic. Chen [27] established the coordination model of main line and entrance ramp based on the ramp inductive algorithm and the variable speed mechanism of the main line of the expressway, combined real-time detection data, and considered factors such as the main line traffic speed, capacity, road demand, and queue space constraints to solve the entrance adjustment rate. Cheng [28] studied the modeling of the main and auxiliary highway equalization control from the local to the overall based on the MFAC model.

2.5 Integration Model of Expressway and Subway Intersections

The integration model of expressway and auxiliary road intersection mainly considers the coordination of system control. The coordination objectives include the timing of signals at adjacent intersections, traffic flow status, and the regulation rate of the entrance ramp. It is an integrated control that aims to maximize the benefits of the expressway mainline while ensuring that the ground road is in good condition. Scholars use different models to coordinate control of expressway and auxiliary road intersections from different perspectives.

Yang [29] linked the traffic flow at ground intersections, ramps, and main roads to determine the minimum average delay as an indicator of the traffic flow status of the ramps, and optimized the entrance regulation rate of the ramps and the signal timing of the adjacent ground intersections. The integrated design of traffic organization and control validates the effectiveness of hierarchical coordination control. Zhou [30] optimized the layered coordination control mode of the adaptation layer, optimization layer, and control layer based on the remaining capacity and on-ramp ratio, with the

objective of reducing the overall congestion level between the main line of the fast lane and the adjacent upstream intersection, effectively reducing system delay time and vehicle queue length. Wang [31] conducted a research and analysis on the unified whole of the three components: the auxiliary road, the ramp, and the expressway. Considering the effect of the traffic flow state at the intersection on the expressway traffic, signal conditioning was performed on a single intersection, and intersections with multiple identical signal cycles were selected. At the same time, secondary road optimization was implemented by using the secondary road as a control target to minimize the overall travel time of the road network system.

Ye [32] established an integrated model of expressway on-ramp and ground intersections, which minimizes vehicle delays as the overall control objective and optimizes the control of inflow rates at the on-ramp, VMS transfer rates, and green letter ratios, and can be effectively controlled queue length at ramp entrance. Kwon et al. [33] adopted the integrated control strategy of the on-ramp and adjacent intersection, which effectively relieved the local congestion and reduced the travel time of the vehicle. Lu [34] considered the influence of signal control at the upstream auxiliary intersection, and analyzed the flow, density, queue length, and vehicle release time at the intersection of the ramp and the expressway with the minimum steady-state error of the system. The MIXED-CONTROL algorithm was improved, proposed a coordinated control model.

3 Review of Control Methods

At present, most of the researches on expressway entrance control are in China. From static development to dynamics, from independent single-point control to system coordination and control, the model is continuously improved. At the same time, with the rapid development of detection technology, data acquisition is rapid and accurate, so that the control parameters of the model can be updated in time to improve the effectiveness and stability of the model. The more in-depth study is the single-channel dynamic control of the entrance of the expressway. The control model based on the density, flow, and occupancy rate is subject to the expressway traffic flow. Because of the non-linear and variability effects, the parameters should be accurately calculated and selected during data smoothing to reduce the errors in the model. Many scholars apply intelligent algorithms to reduce the disturbance on the fast road and improve the application of the model. Although the prediction of the data and the solution of the model are more accurate, it takes a long time to determine the control rules, and the intelligent control model emphasizes theory research, so the specific model needs to be constantly revised in the actual application of the parameters.

For the coordinated control model of expressway entrance, from the perspective of model application, the best indicators of coordinated control are more inclined to maximize the overall efficiency, and determine the maximum traffic volume corresponding to the ramp entrance. From the control theory point of view, that is to

determine the optimal regulation rate of the ramp inlet and to coordinate the control between the inside of the expressway system and the entrance and the adjacent auxiliary road intersection, according to the overall average delay minimum, the shortest travel time for the optimization goal to establish the function to obtain the most excellent solution.

The current coordination control is based on the independent control model and the convergence between the on-ramp links is not smooth enough. When the traffic flow is volatile, the stability of the model is poor, which affects the control effect of the model. When systematically controlling the traffic organization and control, there are many influence parameters and control variables involved, and the calculation is more complex and affects the accuracy of the model. Therefore, after determining the control variables or changing the control mode, it is necessary to calibrate the corresponding parameters according to the actual traffic flow elements after the change, and use the simulation to verify the simulation.

4 Outlook

This paper analyzes the research theories and relates models of the control methods for urban expressway entrance ramps in recent years. Various models use their unique theories and algorithms to optimize the expressway entrances and achieve certain results. The in-depth study of entrance control provides a theoretical basis. In order to make the research on expressway entrance control more applicable to the current development status of expressway and effectively solve the problem of urban expressway entrance congestion, the following suggestions for the entrance control of expressway were proposed:

1. The study on the entrance control of the expressway is to improve its operating efficiency and ensure the high-speed and smooth operation of the main road traffic. Therefore, when building a model, we must consider the composition of the expressway system, from the local depth to the whole, analyze the influencing factors, and realize the coordinated control of the entrance ramp, the main auxiliary road, and the adjacent intersection, so as to promote the overall efficiency of the expressway system. And make the theory more applicable to real-time and variable traffic flow.
2. To realize the coordinated control of the expressway system, it is necessary to analyze the functional components of the road and its influencing factors. The process is more complex and the fluctuation of the traffic flow has a greater impact on data processing. Based on the Internet big data platform, according to the vehicle operating data and the human behavior data, analyze the relevant data of the expressway to find the correlation algorithm between the components, and scientifically solve the adjustment coefficient, so that the established model is more accurate when calculating the entrance adjustment rate.

3. When establishing the inlet control model, it is not confined to the variable selection and determination index of the current model. It should be combined with the actual running status of the expressway and the characteristics of traffic flow, and the integrated control of the expressway and the intersection. In addition to considering the signal timing, the internal influence factors of the intersection should be analyzed, and the evaluation indicators suitable for the established model should be selected so that the coordination effect of the model is more obvious in the application.
4. Consider the connectivity of the regional functions in the process of dynamic coordination and control. When the entry control is implemented, the traffic flow will change accordingly. It is necessary to ensure that the subsequent impact brought by the change of traffic flow is within the control range of the corresponding functional zone. That makes the traffic flow in all areas smooth and improves the continuity and efficiency of coordination control.

5 Conclusion

This paper summarizes the modeling ideas and model characteristics of different models. It analyzes the theoretical basis of the current research on on-ramp control, explored the development process of the model from static to dynamic, from the local to the overall, and described in detail the characteristics of each model. As well as its application effects, the prospects for future research on the entrance control of expressways are also presented. Summarizing the existing models can provide a more in-depth understanding of the modelling method and research ideas for entrance control of the expressway, clarifying the connections and impacts between relevant factors, and laying the foundation for subsequent theoretical research.

Acknowledgements This work is supported by Basic Scientific Research Projects of Higher Education in Liaoning (L17JBKY04) and Social Science Subject of Shenyang (SYSK2019-07-12).

References

1. Wattleworth JA (1965) Peak-period analysis and control of a freeway system
2. Papageorgiou M (1980) A new approach to time-of-day control based on a dynamic freeway traffic model. Transp Res Part B 14(4):349–360
3. Lan YF, Liang XR, Dong CJ (2013) Freeway ramp controllers optimized by ant colony algorithm. Appl Mech Mater 361–363:2214–2218
4. Hu X, Yang J (2009) Study on the ramp control scheme of urban expressway entrance. J Xihua Univ (Nat Sci Ed) 28(04):33–37
5. Liu L, Sun J, Li K (2011) Study on the speed control of the ramp on the entrance of urban expressway. Traffic Inf Saf 29(03):15–19

6. Papageorgiou M, Hadj-Salem H, Blosseville JM (1991) ALINEA: a local feedback control law Foron-ramp metering. Transp Res Rec 1320:58–64
7. Liu X (2013) The intelligent learning control method of ramp on the entrance of express road. Qingdao University of Science & Technology
8. Chi R, Yang X, Li J et al (2013) Fuzzy self tuning PI control method. Control Eng Ramp Entrance Ramp 20(05):813–817
9. Fang C (2016) Study on the flow control of the ramp on the entrance of urban expressway. Southwest Jiao Tong University
10. Yang Z, Pei Y (2002) Study on the fuzzy control of the ramp on the entrance of the expressway. Northeast Highw 25(1):5–8
11. Zhang W, Xiao R, Deng J (2017) Dynamic fuzzy neural network based on genetic algorithm for urban expressway entrance ramp control. Highw Traffic Technol 34(02):129–134 + 148
12. Hu L (2014) Study on the coordinated control strategy of on-ramp of expressway entrance based on MPC. Zhejiang University
13. Cao J (2016) Research on the determination of traffic congestion characteristics and real-time queuing length in urban expressway. Chang 'an University
14. Papageorgiou M, Blosse JM (1990) Modelling and real-time control of traffic flow on the southern part of Boulevard Pdriphdrique in Paris-Part ll: coordinated on-ramp metering. Transp Res Part A 24(5):361–370
15. Gao X (2017) Research on the optimization of ramp control at the entrance of the urban elevated expressway based on ALINEA. Southwest Jiaotong University
16. Feng C, Jia Y, Jian L et al (2011) Design of fuzzy neural network control method for ramp metering. In: Third international conference on measuring technology and mechatronics automation. IEEE Computer Society, pp 966–969
17. Li HF, Wang XF, Ding QY (2012) Switching control methods for expressway Ramp metering based on hybrid automaton model. Adv Mater Res 591:1346–1350
18. Fares A, Gomaa W (2015) Multi-agent reinforcement learning control for Ramp metering. Progress in systems engineering. Springer International Publishing, 167–173
19. Liu P, Ren J (2011) The expressway entrance ramp control. Sci Technol Eng Based on PID Neural Netw 11(32):7961–7966
20. Zhu H, Liu X, Shi Q (2010) Expressway entrance control based on task balancing model. J Transp Syst Eng Inf Technol 10(5):85–90
21. Zhu H, Liu X, Shi Q (2010) The control model and application of the expressway entrance control based on the balanced thought. Transp Syst Eng Inf 10(05):85–90
22. Lu XY, Varaiya P, Horowitz R. Novel freeway traffic control with variable speed limit and coordinated Ramp metering. Transp Res Rec 2229(2229):55–65
23. Sun J, Zhang S, Tang KS (2014) Online evaluation of an integrated control strategy at on-ramp bottleneck for urban expressways in Shanghai. IET Intell Transp Syst 8(8):648–654
24. Zhu J (2012) Research and implementation of variable speed limit control method for expressway entrance ramp and road section. Zhejiang University of Technology, Hangzhou 32–46
25. Liang Z (2016) Study on the strategy of the coordinated control strategy of the entrance ramp of the expressway and the variable speed limit. Harbin Institute of Technology
26. Yang Y (2009) Study on the coordinated control of the entrance ramp and road section of expressway. Chongqing Transp Sci 04
27. Chen J, Guo J, Lu L et al (2013) Study on the coordinated control strategy of fast speed under variable speed control. J East China Jiaotong Univ 30(02):47–51
28. Cheng Z (2015) Balanced control of main and auxiliary road system based on MFAC. Beijing Jiaotong University
29. Yang H (2012) Research on the integration of traffic organization and control of urban expressway. Kunming University of Science and Technology
30. Zhou Q (2017) Study on the coordinated control method of the ramp of the expressway entrance and the interrelated intersection. Beijing Jiaotong University

31. Wang X (2017) Study on the entrance ramp control of urban expressway. North China University of Technology
32. Ye H (2014) Research on coordinated control of ramp and ground intersection of urban expressway entrance. Kunming University of Science and Technology
33. Kwon E, Ambadipudi RP, Bieniek J (2003) Adaptive coordination of ramp meter and intersection signal for optimal management of freeway corridor
34. Lu Z (2013) Research on the coordinated control algorithm of the ramp and auxiliary road signal of urban expressway. North China University of Technology

Numerical Study on the Effect of Driving Distance on the Diffusion of PM2.5 in the Street

Peng Xu, Mengru Wang, Xi Lu, Junru Han, Qin Gu and Chen Ma

Abstract The influence of inter-vehicle distance on the diffusion process of PM2.5 in the underlay surface of urban streets is studied. Firstly, the diffusion process of tracer gas PM2.5 is simulated using CFD, and the emission factors of PM2.5 under different vehicle types and speeds is analyzed and summarized. Then, using the improved MIRA vehicle model, a PM2.5 diffusion model of the traveling vehicle at different intervals was established. In the end, POST post-processing was applied to the model, and the PM2.5 concentration field cloud map and the velocity field cloud map around the vehicle body were compared and analyzed to obtain the diffusion rule of PM2.5 emission from the vehicle exhaust gas.

Keywords Driving distance · PM2.5 · Numerical simulation · Tracer gas

1 Introduction

More than half of the city's air pollution comes from exhaust emissions from motor vehicles that use gasoline and diesel as fuel [1]. In the streets and roads where a large number of motor vehicles exist, the total amount of PM2.5 emitted is large [2]. At the same time, the traffic wind caused by the large number of motor vehicles in the street result that the PM2.5 ultrafine particles at the bottom of the street will spread to the street, affecting pedestrians [3]. Therefore, it is of great significance to study the diffusion process of motor vehicle emissions under the influence of traffic factors.

P. Xu · M. Wang · J. Han · Q. Gu
College of Civil and Transportation Engineering, Hohai University,
Nanjing 210098, Jiangsu, China
e-mail: pxu5@iit.edu

X. Lu (✉)
China Urban Sustainable Transport Research Center
China Academy of Transportation Science, Beijing 100029, China
e-mail: allenlu87@126.com

C. Ma
School of Astronautics, Harbin Institute of Technology (HIT), Harbin 150001, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_24

The proliferation of vehicle exhaust PM2.5 is affected by various factors such as the spatial form of the underlying surface and vehicles. Huang et al. [4] studied the dimensionless pollutant concentration distribution on the windward and leeward sides of the canyon and compared the simulation results. The canyon model is taken as the spatial pattern of urban streets, the movement of vehicles becomes an active factor affecting the diffusion of vehicle emissions PM2.5. Zhang [5] researched the numerical analysis and optimization of the effect of the body's tail structure on the wake flow field of the car. Xiucheng Li and Haipeng Liu's articles [6] considered the impact of the single-car models and tail fins on the wake. He and Fu [7] studied the influence of factors such as vehicle shape, vertical spacing, and traffic volume on the external flow field of queued vehicles. The research on the diffusion process of PM2.5 in the street mainly focuses on the wake structure of the single car and the wake diffusion between the teams. However, the microscopic scenes caused by the workshop airflow pattern and the diffusion mechanism of PM2.5 in the workshop have not been studied in depth, such as the vehicles in the road, due to the mutual movement status. At the same time, under different vehicle speeds and conditions, the numerical difference between the vehicle PM2.5 emission factors is greater [8].

The use of secondary aerosol generation simulations and fine particle diffusion studies are considered. Through on-site investigation, mathematical derivation, model correction and other methods, the diffusion characteristics of exhaust gas under different driving spacing conditions are analyzed. Firstly, the CFD numerical simulation is used for the diffusion process of the tracer gas PM2.5. The emission factors of PM2.5 under different vehicle speeds for different models are analyzed and summarized. Improved MIPA model is used combined with fluent software. A CFD model is established through analyzing the background conditions and determining the boundary conditions and grid scheme. POST processing is used to compare and analyze speed vector and concentration cloud maps in the three directions of motor vehicles x, y and z in order to obtain the diffusion rule of vehicle exhaust gas emissions PM2.5.

2 The Assumptions and Analysis of the Influence of Inter-vehicle Distance on PM2.5 Diffusion

2.1 The Diffusion Mechanism of PM2.5

During the three states of motor vehicles intersecting, disjointing and following-up, due to the setting of infrastructures such as the central separation belt, the motor vehicles are greatly affected by the preceding vehicle in the car-following state. The diffusion characteristics of PM2.5 emissions from motor vehicles in the car-following state is studied. The driving speed of the vehicle largely affects the PM2.5 diffusion of motor vehicle exhaust emissions [9, 10]. Due to the largest proportion of cars in urban road, the selection of cars as the model is the most representative. The speed of

urban roads is not high, especially in the state of congestion at the peak, the exhaust emissions of motor vehicles are the most serious [11] and they are considered as representative research scenarios. Therefore the PM2.5 diffusion model for urban roads at low speeds is established. The vehicle speed is 5 km/h.

The CFD numerical simulation is used to study the diffusion characteristics of tail gas in different wake fields. The content is as follows. The effect of background concentration on the road, crosswind influences, and heat exchange isn't considered. A single vehicle emission model for motor vehicles is established. The boundary value is set. The right computing domain and grid solution is chosen. To study the diffusion characteristics of tracer gas in the wake field of a standard vehicle model when the distance between vehicles is changed. PM2.5 is selected as tracer gas. The fast-back vehicle model in the MIRA standard model is used to calculate vehicle emissions.

2.2 Determination of Tracer Gases and Emission Factors

PM2.5 was added as a tracer gas to the model to study the diffusion of PM2.5 in different distances of the vehicle flow field. Due to its unique conditions, street canyons have lower gas flow rates in the streets. Therefore, it is considered as an incompressible fluid at the time of calculation. The MIRA model is a research object for wind tunnel tests and does not have the ability to discharge itself. Therefore, it is necessary to select a suitable vehicle PM2.5 emission factor as a data source to add to the vehicle exhaust diffusion model.

PM2.5 has a complex composition and different combinations of PM2.5 sources. The emission of PM2.5 from motor vehicles will undergo complex chemical and physical reactions during the diffusion process. There are no accurate conclusions and public expressions on these chemical and physical reactions. Therefore, it is assumed that the chemical properties of PM2.5 emissions are stable and do not change with temperature and speed. Previous researches on the ultra-fine particulate matter in the wake field of motor vehicles have tended to directly specify the total emission of motor vehicle PM2.5. The consideration of the vehicle's PM2.5 emission factor is slightly insufficient. Combined with the PM2.5 emission factor, the tracer gas is defined more detailedly. At present, bench test method, tunnel test method and induction test method are mainly used to study PM2.5 emission factors [12]. Tunnel test method can obtain comprehensive emission factors of vehicles in real road sections. Emission factor obtained by tunnel test method is selected as data source for PM2.5 tracer gas. The emission factors of different models of PM2.5 under different speeds obtained by the research team are shown in Table 1.

The PM2.5 diffusion model for urban roads at low speeds is established. The speed of the vehicle is 5 km/h, and the emission factor of the corresponding car is 0.054 g vehicle/km. The exhaust emissions of motor vehicles and the composition of tail gas have a lot to do with the driving speed of motor vehicles, the load of motor vehicles themselves, and the driving environment. According to the relationship [13]

Table 1 PM2.5 emission factors of urban motor vehicles (g vehicle/km)

Speed (km/h)	Taxi (natural gas)	Car (gasoline)	Bus (natural gas)	Coach (Diesel)	Truck	Average error
5–20	0.01871	0.05351	0.09108	0.8012	0.1691	0.29
20–35	0.01851	0.05312	0.08433	0.7543	0.1643	0.31
35–40	0.01762	0.05245	0.07209	0.7065	0.1597	0.31
40–45	0.01522	0.05230	0.06693	0.6496	0.1465	0.30
45–50	0.01521	0.04895	0.06427	0.6234	0.1401	0.33

between different driving speed and exhaust emissions the PM2.5 emission of the car is 5.8×10^{-4} g/s after conversion.

MIRA car models that are widely used in the research community at home and abroad for research. The MIRA model has a simple structure and is relatively simple when establishing a motor vehicle model, which is conducive to modeling and analysis. This simulation selected the fast back model in the MIRA standard model. It is calculated that when the vehicle speed is 13 m/s, the displacement of the MIRA model is 9.6×10^{-3} g/s according to the size and formula [14] of the fast back vehicle model [15].

When defining the PM2.5 gas emitted by motor vehicles, the use of tracer gas (PM2.5) method is considered to study the diffusion characteristics of PM2.5. Therefore, the changes in the physical and chemical properties of other gases in the exhaust gas and the exhaust gas is ignored, and the amount of PM2.5 only is considered. According to the calculation, the mixture of PM2.5 emissions from motor vehicles is defined in Fluent, with PM2.5 accounting for 5.8% and other gases accounting for 94.2%.

3 PM2.5 Diffusion Numerical Models for a Given Vehicle Speed and Model

3.1 Establishing a CFD Model for Vehicle Emissions

When domestic and foreign scholars study the pollution process of motor vehicle exhaust diffusion, they generally use the following three methods: wind tunnel test [16], field test, and CFD numerical simulation. The core of the proliferation problem of motor vehicle exhaust emission PM2.5 is also the turbulence problem. The turbulence problem is just a typical fluid mechanics problem. The process can be analyzed by CFD (computational fluid dynamics) [16]. Jean et al. established a numerical model to simulate the concentration of pollutants in asymmetrical street canyons and explained the relationship between the concentration distribution of model results and the street canyons. The MIRA model used is easy to construct, simplifying a right

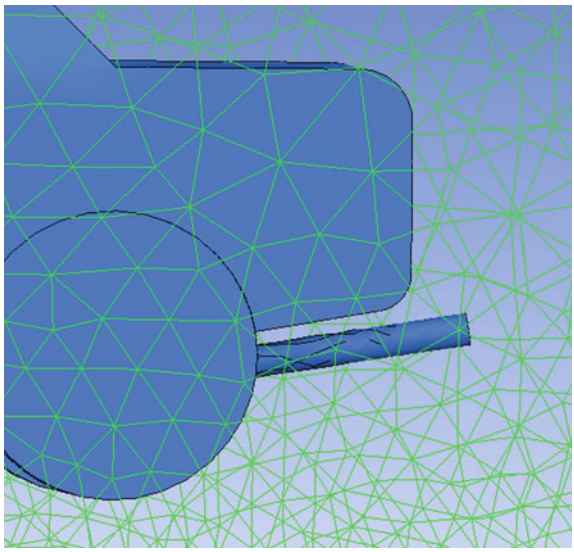


Fig. 1 Exhaust pipe grid

angle at the front end to the fillet, reducing the amount of calculation and modeling difficulty, and the simulation results are not much different from the results of wind tunnel tests and domestic and foreign researchers. In order to study the diffusion process of motor vehicle emissions PM2.5 in the flow field outside the vehicle, it is necessary to add tracer gas. The exhaust pipe is located 35 cm to the left of the axis of the vehicle, 10 cm beyond the tail of the vehicle, and the exhaust pipe is 4 cm in radius. In the study of the diffusion of motor vehicle emissions PM2.5, in order to consider the effect of heat transfer on the diffusion process, the exhaust pipe is lifted upwards by 9°. When using Fluent modeling, the grid is stretched at the rear of the car according to the size of the exhaust pipe to form a surface as the quality exit of the exhaust pipe of the motor vehicle. Exhaust pipe position and grid are as follows (Fig. 1).

3.2 Computing Domain and Grid Scheme

The coordinate system in the Fluent vehicle model is set according to the following rules:

(1) The direction of the front of the vehicle to the rear of the vehicle is the positive direction of the X-axis, and the front of the first vehicle is $x = 0$.
(2) The vertical upward direction is the positive direction of the Z axis, and the horizontal plane of the exhaust pipe is $z = 0$.
(3) According to the right-hand rule, the positive direction of the Y axis is determined, the left edge of the body is $y = 0$.

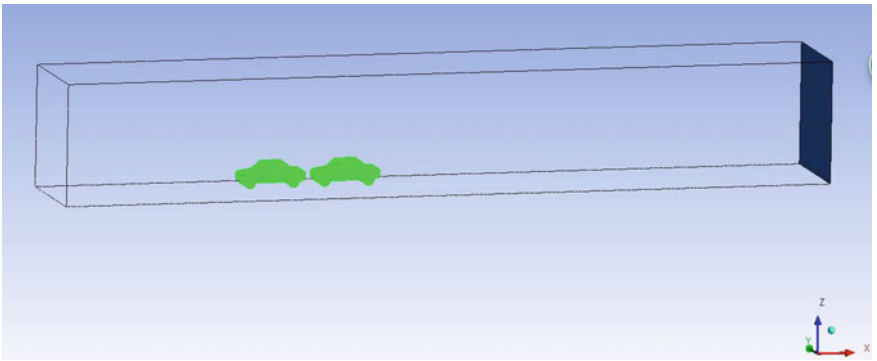


Fig. 2 Car diffusion numerical simulation model coordinate system

(4) Considering the speed and the calculation ability comprehensively, the rear plate of the first car is selected to be 800 mm and 1600 mm from the front of the rear car respectively. The coordinate system is shown in Fig. 2.

When performing CFD numerical simulation, it is necessary to consider whether the established computational domain will have a blocking effect and how to make the blocking effect negligible. According to the research [17], when the blocking ratio is between 1 and 2.5%, the interference errors due to obstruction can be ignored, The blocking effect is calculated as follows.

$$\varepsilon = \frac{A_1}{A_{N1}} \tag{1}$$

Among them, A_1 represents the orthographic projection area of the model, where the value is 1.856 m^2, and A_{N1} represents the import area of the calculation domain. The value here is $7.605 \times 11.375 = 86.51$ m^2, and the obstruction effect value is calculated as $2.14\% < 2.5\%$. Meet the requirements of the blocking ratio, the blocking ratio when performing numerical calculations is considered.

(1) In order to ensure the computational efficiency and accuracy of the model, the computational domain of the numerical simulation model is as follows.
(2) The entrance to the MIRA model is three times the length of the front section of the MIRA model, which is 12495 mm. The distance between the first car and the second car is 800 and 1600 mm, and the length of the second car is 7 times the length of the car, which is 29,155 mm.
(3) The top of the model is still 4 times the height of the roof, which is 6084 mm.
(4) The model is still about three times the width of the car, which is 4875 mm.

In the process of using the hybrid grid scheme, it was found that for the external flow field of the vehicle MIRA model, it was noted that the flow conditions around the vehicle were more complicated, especially the angle between the front window and the engine compartment and the corners of the engine compartment, and the exhaust

tube is located at the tail of the MIRA model, and PM2.5 exhaust gas is more affected by the exhaust gas field. In order to meet the requirements of calculation accuracy and computer computing ability, it also meets the requirement of accurately reflecting the details of the flow field in the rear of the motor vehicle. The MIRA model is encrypted at the front corners of the car and at the corners of the engine room, at the corners of the car body, and at the rear of the vehicle. The specific encryption processing method is: the body surface grid size is 80 mm; the triangular prism grid is stretched, the first grid interval is 80 mm, and the grid separation number is 30. Figure 3 shows the MIRA model encryption process when the vehicle spacing is 1600 mm.

Figure 4 The distance between driving distances is a calculation map of a 0.8 m-range vehicle calculation area, which ensures that the boundary conditions of the vehicle models with different distances are the same, only changing the distance between vehicles, and comparing the influence of driving distance on the PM2.5 diffusion process of motor vehicle emissions. Therefore, the boundary conditions in the three different spacing vehicle models are as follows (Table 2).

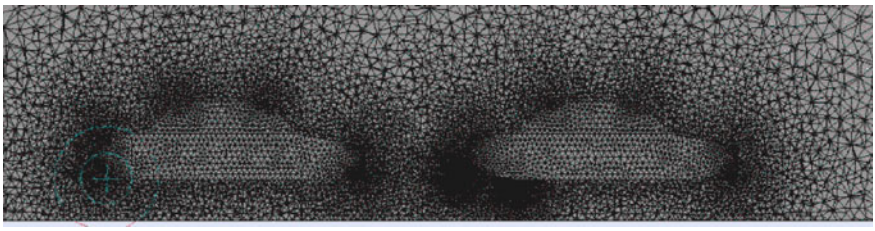


Fig. 3 MIRA model encryption map (vehicle spacing is 1600 mm)

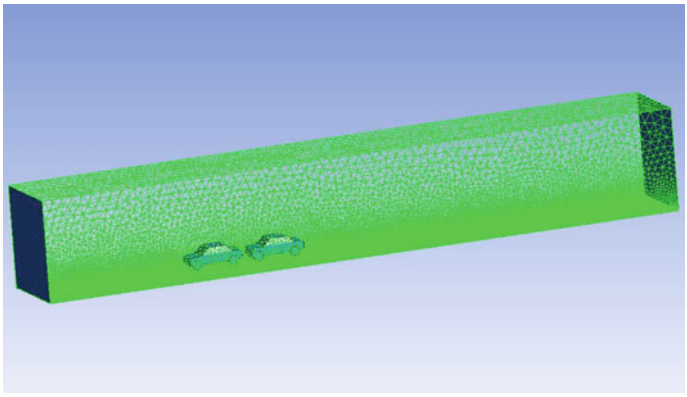


Fig. 4 Rendering of vehicle calculation domain with 0.8 m driving distance

Table 2 Numerical model boundary conditions for vehicle emissions

Boundary location	Boundary conditions	Settings
Calculation domain front end	Entrance	Speed entrance, v = 13 m/s
Computing domain backend	Export	Pressure outlet, static pressure = 0 Pa
Left rear end of vehicle model	Exhaust pipe inlet	Speed entrance, V = 1 m/s (94.2% of the air in the exhaust)
Body surface, calculation area ground	Wall surface	Non-slip wall
Top of the calculation domain	Wall surface	Non-slip wall
Both sides	Wall surface	Non-slip wall

4 Comparison and Analysis of Results

Post-post processing is used to extract the speed vector with image types. The results of numerical simulation are as follows.

The concentration distribution of PM2.5 in the x = 4.8 m section of the motor vehicle model under different vehicle distances (from left to right, single-car, 0.8, and 1.6 m). It can be seen from Fig. 5 that the PM2.5 gas emitted by the motor vehicle has risen. When the bicycle is traveling, the height of the lift is low, so as to better observe the case of diffusion of PM2.5 exhaust gas at the rear of the vehicle take the speed cloud map with x = 4.8 m section. Figure 6 shows the speed cloud diagram

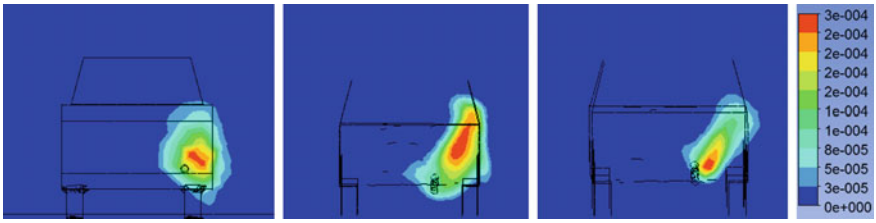


Fig. 5 Plot of PM2.5 concentration in 1x = 4.8 m section (vehicle, 0.8, 1.6 m)

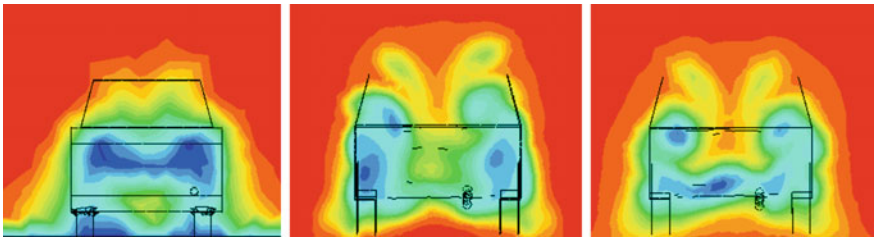


Fig. 6 The speed profile of the x = 4.8 m section (bicycle, 0.8, 1.6 m)

Fig. 7 PM2.5 concentration cloud diagram of y = 0.8125 m section (bicycle, 0.8, 1.6 m)

for different vehicle distances. It can be seen from the figure that the velocity vector is symmetrical about the center axis along the central axis of the vehicle's MIRA model. Characteristically, there are symmetrical vortices near the lower end of the tailgate of the motor vehicle, which is consistent with the results shown by the cloud map of the concentration field.

Combining the velocity cloud map with the PM2.5 concentration cloud map, it can also be analyzed that the low speed airflow around the motor vehicle is mainly distributed on the left and right side of the vehicle body and the tail of the vehicle body, while the PM2.5 exhaust gas diffusion process mainly occurs in the tail region of the vehicle, with only a few gas crosses the side panels of the vehicle body and escapes the rear wake area of the vehicle. This is because there is a relatively high-speed airflow on the side of the vehicle body.

Figure 7 shows the concentration cloud of PM2.5 in the section y = 0.8125 m. It can be seen from the figure that the PM2.5 diffusion distance is long when the vehicle is driving when the vehicle is in one -way driving. When the car is driving with the car, It has become relatively short. This is due to the obstruction of the MIRA motor model, which makes the PM2.5 discharged from the motor vehicle unable to flow freely. The distribution of the concentration field at the distances of 0.8 and 1.6 m is also quite different. When the distance is 0.8 m, due to the short distance from the front car, the rear vehicle field is more complicated. Under the influence of the tail vortex, the PM2.5 exhaust gas diffuses from the second roof and the bottom of the vehicle at the same time, but the diffusion amount is relatively low. As a result, the concentration of PM2.5 in the second vehicle is low that can be found on the axial surface of the vehicle, and when the distance between vehicles is 1.6 m, Due to the relatively long distance between the two vehicles, the PM2.5 exhausted from the vehicle in wake area is fully diffused and is pulled by the vehicle at a higher speed. It flows rearward through the bottom of the vehicle. Therefore, it can be seen from the figure that a certain concentration of PM2.5 gas exists behind the second vehicle.

In order to better understand the diffusion of PM2.5 around the vehicle body, the PM2.5 concentration cloud map and velocity cloud map at z = 0.6 m (plane where the vortex center is located) are taken. As can be seen from the figure, when the vehicle is in one -way driving, the spread of PM2.5 exhaust is more concentrated, compared to the case where the distance between the bicycle and the running distance is 0.8 m, and the distance behind the vehicle is 1.6 m. The range of motor vehicle wake fields is smaller, but in the vicinity of the vehicle, the influence of wake field has a larger range, which affects the spread of PM2.5 (Fig. 8).

Figure 9 is a speed cloud diagram of the movement of PM2.5 gas outside the vehicle after the distance between the two vehicles is 0.8 and 1.6 m: laterally comparing the three diagrams, it can be clearly seen that when a single vehicle is driving, the main diffusion area of PM2.5 is in the tail wake area of the vehicle. PM2.5 exhaust gas will not substantially exceed the outer contour of the vehicle during the diffusion process. At the same time, when the PM2.5 exhaust gas diffuses in the space, it will appear outward as Spiral flow. When the distance between vehicles is 0.8 and 1.6 m,

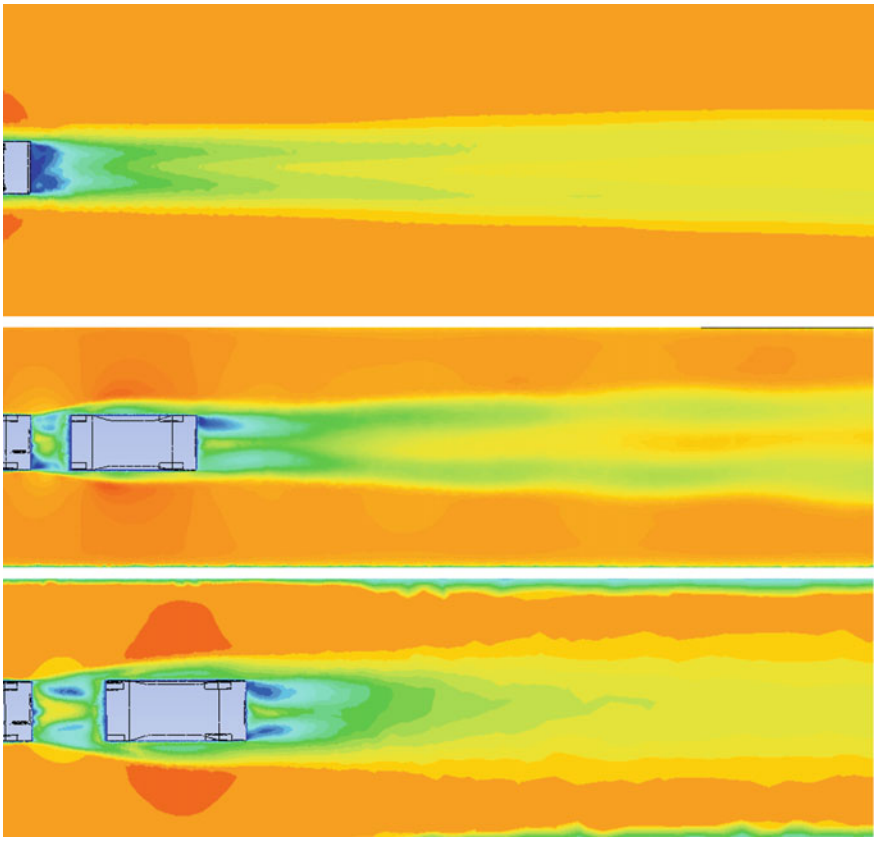


Fig. 8 The plot of PM2.5 concentration in the z = 0.6 m section (vehicle, 0.8, 1.6 m)

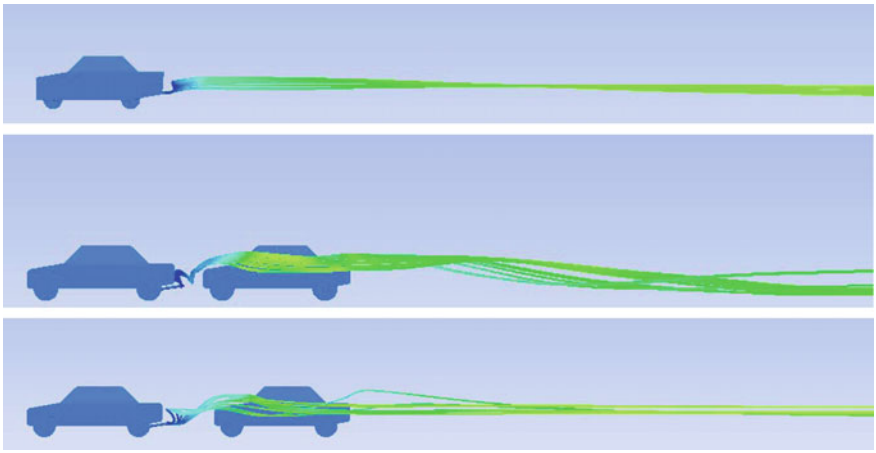


Fig. 9 Speed streamline diagram of motor vehicle PM2.5 exhaust (bicycle, 0.8, 1.6 m)

the PM2.5 exhaust gas also exhibits a spiral structure, but when the PM2.5 exhaust gas spreads out of the vehicle, by passing the rear-car driving vehicle and moving far away, spreads distantly and spreads longer, the area with greater diffusion impact is the motor vehicle's tail and the periphery of the second vehicle body, and it dissipates faster after the second vehicle.

In the final analysis, the influence of the distance between motor vehicles on the diffusion process of PM2.5 is mainly reflected in the influence of the distance between the driving and the wake field. When the distance between vehicles is small, the wake field is more complex between vehicles and the tail vortex cannot be detached in time. Therefore, PM2.5 will not only disperse at the tail but also will rise upward when the tail flow field influences; At the time, the impact of the rear vehicle on the front vehicle wake flow is reduced, and the rear vehicle is more used to displace the front air and cause the PM2.5 exhaust gas to spread to the periphery. Therefore, when the distance is 1.6 m, the PM2.5 exhaust around the vehicle will be more. It can be seen that the driving distance of motor vehicles has a significant impact on the diffusion process of PM2.5.

5 Conclusion

Using the MIRA vehicle model, an exhaust pipe was first added to the tail of the model, and PM2.5 was used as a tracer gas. Then, a model was established to simulate diffusion of exhaust gas PM2.5 under three conditions: single-vehicle traveling, vehicle distance as 0.8 m, and vehicle distance as 1.6 m. The POST post-processing software was used to process the obtained results. The effects of the

distance between vehicles on the diffusion of exhaust gas PM2.5 was studied by comparing the diffusion conditions under the three conditions.

Studies have shown that under different driving distance conditions, PM2.5 diffusion presenting different rules. When single-vehicle traveling, PM2.5 diffusion is mainly distributed in the rear of the vehicle, and the diffusion range is wider; when the distance between two vehicles is 0.8 m, Due to the influence of the rear vehicle, the structure of the front tail flow field is complex. Under the influence of the wake flow field, the PM2.5 diffusion process mainly occurs between the front and rear car bodies, Only a little PM2.5 gas diffused across the roof under the influence of turbulence in the wake zone; when the distance between the two vehicles is 1.6 m, the impact of the rear vehicle on the flow of the front tail is gradually reduced, PM2.5 under the influence of high-speed airflow around the bottom and the body of the vehicle, mainly diffused to the left side of the body and the bottom of the vehicle, and the diffusion range is wider.

Acknowledgements Supported by the natural science foundation of Jiangsu: (No. BK20151497).

References

1. Zhang X (2014) PM2.5 generation source and countermeasures. Examination Weekly, vol A5, p 194
2. Zhou Y (2016) Establishment of a high-resolution emission inventory and its evaluation through air quality modeling for Jiangsu Province, China. Nanjing University
3. Zhou Z (2013) Study on the convective diffusion law of PM2.5 emission from vehicle exhaust in urban street canyon. Shandong University of Architecture
4. Huang Y, Wang S, Jin X et al (2008) A comparative study of various turbulence models for simulating pollutant dispersion within an urban street canyon. Hydrodyn Res Prog A 02:189–195
5. Zhang J (2015) Study on the effects of body rear structure to wake field and the body rear structure optimization by computational fluid dynamics. Chongqing Jiaotong University
6. Li X (2014) Numerical simulation of wake flow field effect on exhaust dispersion characteristics of a car. Jilin University
7. He B (2009) Research on automotive aerodynamic characteristic of platoon. Jilin University
8. Li Y (2014) Investigation of exhaust PM 2.5 emission profile of typical motor vehicles. China Academy of Environmental Sciences. Jilin University
9. He L, Hu J, Zu L et al (2013) 2015 emission characteristics of exhaust PM2.5 and its carbonaceous components from China I to China III heavy-duty diesel vehicles. Acta Sci Circumstantiae 35(3):6.56–662
10. Beevers SD, Carslaw DC (2005) The impact of congestion charging on vehicle speed and its implications for assessing vehicle emissions. Atmos Environ 39(36):6875–6884
11. Tian L, Fan S, Zhang D et al (2016) Influent of average speed on vehicle exhaust emissions. Environ Eng 11:6541–6548
12. Liu J (2009) The distribution law of motor vehicle exhaust in the street valley and its impact on the built environment. Harbin Institute of Technology
13. Wang B, Zhang Y, Zhu C et al (2001) Research on urban vehicle emission factors tunnel test. Environ Sci 02:55–59
14. Li Y (2015) Simulation study on PM2.5 emission and dispersion of automobile tail. Chang'an University

15. Kubilay A, Neophytou MKA, Matsentides S et al (2017) The pollutant removal capacity of urban street canyons as quantified by the pollutant exchange velocity. Urban Clim 21:136–153
16. Deng C, Bang S (2014) ANSYS FLUENT 14.0 simulation analysis and optimization design. Sun Machinery Press
17. Nozu TTT (2012) Les of turbulent wind and gas dispersion in a city. J Wind Eng Ind Aerodyn 104–106(3):492–499

Study on the Speed Limit of Vehicle Stability Under Rainy Environment

Peng Xu, Kai Jiang, Xi Lu, Junru Han, Chen Ma and Xinran Xu

Abstract The study on the lower limit of adverse weather conditions in China's current expressways is mostly considered in terms of road conditions and traffic flow, it is mostly based on the parking distance based on speed control research and does not take into account the stability of the vehicle in bad weather. This paper studies the stability of vehicle driving under rainfall conditions and obtains safe driving speed through ADAMS/Car simulation simulation, it provides a reference for the system's comprehensive establishment of the highest safe driving speed of the highway in the rainy environment.

Keywords Highway · Rainy weather · Stability analysis · Vehicle speed limit · ADAMS/car

1 Introduction

Road slippery in the rain affects vehicle stability, when the vehicle is driving at high speed, it has a serious impact on the safety of the vehicle. In foreign countries, Sherretz analyzes driving safety in rainy weather based on real weather data, the study considers that there is a negative linear relationship between traffic safety and rainfall intensity [1]. Soyoung considers the impact of rainfall on driving safety from the perspective of safe driving of individual vehicles [2]. Schwab conducts research on driving safety in inclement weather and finds that the speed of the highway on rainy days is significantly reduced [3]. Hall study believes that the speed of vehicles under light rain conditions is basically unaffected, the speed of vehicles in

P. Xu · K. Jiang · J. Han
School of Civil and Transportation, HoHai University, Nanjing 210098, China

X. Lu (✉)
China Urban Sustainable Transport Research Center,
China Academy of Transportation Science, Beijing 100029, China
e-mail: allenlu87@126.com

C. Ma · X. Xu
School of Astronautics, Harbin Institute of Technology (HIT), Harbin 150001, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_25

heavy rain drops significantly by 3 to 6 mil/h [4], In the domestic aspect, domestic experts and scholars started late on the impact of traffic safety in adverse weather conditions, the impact of rainy weather on traffic safety is mainly discussed in terms of traffic flow, speed and distance. Yin Tao [5] and others consider the road adhesion coefficient in part of the waterslide, constraints on visibility under different rainfall conditions, studying safe speed values in rainy conditions. Ji et al. [6] analyzed from the perspective of side-slip accidents on rainy days, obtained water film thickness models under different rainfall conditions and determine the safe driving speed of the vehicle. Cheng Guozhu [7] analyzes the mechanism of rain weather conditions on road traffic safety, limiting the speed of vehicle safety on the freeway based on the parking distance model and proposed the speed limit standards for different visibility distances, adhesion coefficients and longitudinal slopes for rainy days. The research on the impact of speed restrictions on road speeds under the conditions of rainy days at home and abroad is mostly based on historical data of road accidents, this method is limited by objective conditions, and historical data lacks timeliness, it is also impossible to analyze the safety speed of vehicles under rainfall conditions from a fundamental or stability point of view. Therefore, this paper considers the analysis of the stability of the vehicle under rainy conditions, it is found that the rainfall environment influences the driving environment, road friction coefficient and braking distance through the analysis of the mechanism of rainfall disasters. Finally, using ADAMS/Car simulation to obtain a safe driving speed, it provides a basis for reasonable speed limit under bad weather conditions.

2 Vehicle Stability

2.1 Anti-sliding Stability Analysis of Vehicles

The tire side slip angle is a key index to characterize the yaw stability of the vehicle, the slip angle is the angle between the motion trajectory of the front and rear wheels of the vehicle and the forward direction of the wheel coordinate system when the vehicle is running or braking [8]. It directly reflects the ability of the vehicle to follow the desired trajectory, it is often used to analyze, control, and evaluate the stability characteristics of a vehicle. The slip angle α is related to the lateral force F_y, as shown in Fig. 1.

The main reason for the loss of yaw stability of a vehicle is the non-linear lateral deflection of the tire. When changing within a certain range, F_y will have a nonlinear increasing relationship with α, Once α exceeds the threshold range and F_y saturated, explain that this time F_y reaches the limit, the yaw moment of the vehicle gradually decreases toward zero andthe vehicle cannot maintain lateral stability. The travel speed corresponding to the limit state of tire F_y is the critical speed of the vehicle, when the friction coefficient of car surface becomes smaller, the critical speed also decreases gradually. In the rainy day, the vehicle traveling on the road surface has a

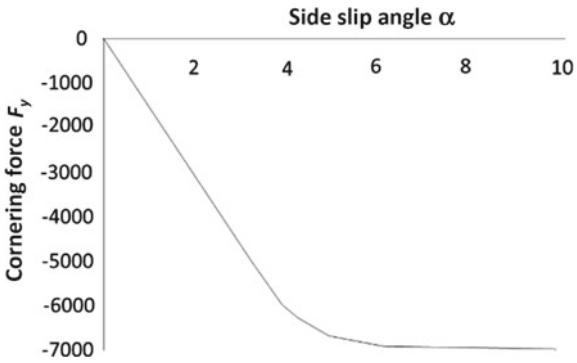


Fig. 1 Tire deflection characteristics

water film, and the road surface adhesion coefficient decreases, making the vehicle braking distance longer. The wheel adhesion coefficient is asymmetrical, and the wheel load and road adhesion are also reduced, when the vehicle is over-steering at the corner, once there is a slight lateral force, it will cause the vehicle to skid.

The lateral adhesion coefficient is the ratio of the lateral force to the surface normal reaction force [9]. According to $F_y = mg\varphi_y$, F_y related to lateral force, it can be seen from Fig. 2 that the lateral force is related to the side slip angle. Therefore, the lateral adhesion coefficient must be related to the side slip angle. The lateral adhesion coefficient φ_y is also related to the tire slip rate and slip rate, as shown in Fig. 3.

Analyze the curve, when the vehicle is in the driving state, the lateral adhesion coefficient decreases sharply with the increase of slip rate; when the vehicle is in a braking state, the lateral adhesion coefficient decreases sharply with the increase of slip rate.

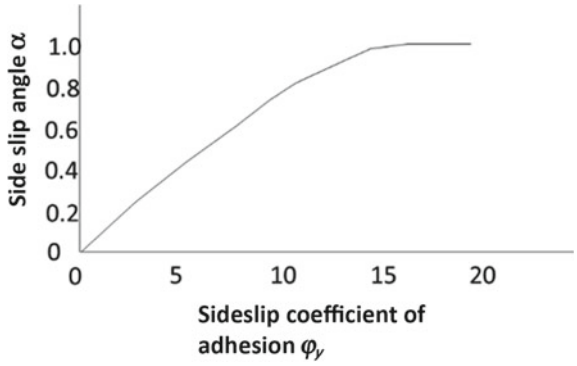


Fig. 2 Tire deflection characteristics

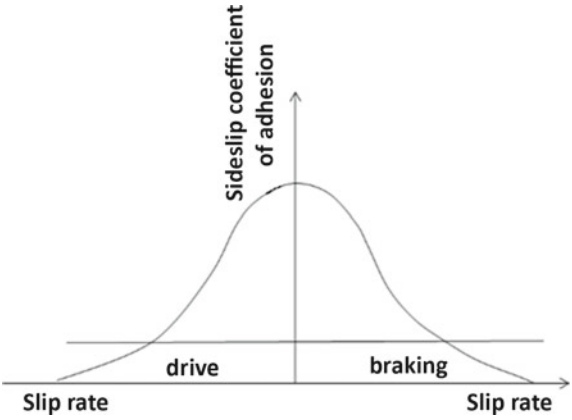


Fig. 3 The relationship between lateral adhesion coefficient and slip ratio and slip rate

2.2 Anti-rollover Analysis of Vehicles

The hazard of rollover is far greater than the risk of a rollover accident, lateral slip generated by the driver when controlling the steering wheel angle of the vehicle, once the vehicle hits a road or an obstacle, a tripping rollover accident occurs. This dangerous condition is not controllable, so this article does not analyze it [10]. When the vehicle is not affected by obstacles, it is affected only by the lateral stability, and a non-flooding rollover accident occurs. Due to the excessive steering operation of the driver, the side weight of the lateral acceleration tire of the vehicle can be compensated by a limit value, which causes the vertical load of one side tire to decrease to zero, and the vehicle will roll over.

2.3 Analysis of Water Slippage in Rainy Day

After the rainfall, the water on the road will form a water film on the road surface. When the thickness of the water film reaches a certain level, as the vehicle speed increases to a certain value, the tire and the road surface will be disengaged, the adhesion coefficient is basically zero, and the phenomenon of complete water-skipping occurs, corresponding driving speed is critical speed.

When a high-speed vehicle is driving on a surface of accumulated water, the adhesion coefficient of the road becomes smaller, and the vehicle is liable to slip. In addition, when the tire is in contact with rain, rainwater will enter the tire and the vehicle brake, weakening the braking performance of the vehicle. The vehicle is traveling on the road. Because the driving wheel and the braking wheel cannot drive on the water road at the same time, the driving force and braking force of the left and right wheels are different, when the driving state of the vehicle changes, the vehicle easily slips to the side of the tire running on the surface of the water. In rainy

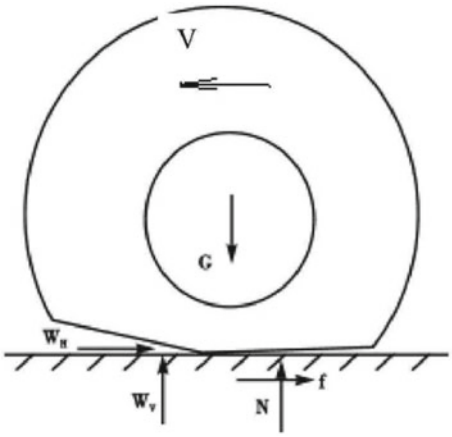


Fig. 4 Analysis of the force when the tire is skiing

conditions, the resistance created by the water between the tire and the road causes the tire to float, causing a "slippery" phenomenon in the event of danger. In extreme conditions, the car loses control. When the tire is skiing, the stress situation is shown in Fig. 4.

As shown in Fig. 4, as the wheel turns, the rain on the road surface will gather under the rotating tire and form a wedge shape, the wedge-shaped water film is driven by the tire and has a backward movement. At this time, the water film generates an upward force W_V and a backward force W_H on the tire. As the water film gradually gathers, the hydrodynamic pressure here also gradually increases, when the hydrodynamic pressure exceeds the wheel pressure, the area of the contact surface between the tire and the road surface continuously decreases. When the contact area is reduced to 0, the vehicle braking is limited, resulting in water slip phenomenon.

3 Instability Model Construction

3.1 Anti-skid Stability Model

In this study, the analysis of the turning of the side-sliding motion in response to the limit of turning conditions considers the lateral acceleration and the yaw rate of the vehicle. The force when the vehicle makes a turn is shown in Fig. 5.

The turning circular motion mechanics analysis of Fig. 5 shows that the centrifugal force of the vehicle is $F_c = \frac{mu^2}{R}$; The lateral adhesion resulting from the contact of the wheel with the ground is $F_y = mg\varphi_y$. To ensure that the vehicle does not slip when turning, must meet the conditions $F_y \geq F_c$:

$$\frac{mu^2}{R} \leq mg\varphi_y \tag{1}$$

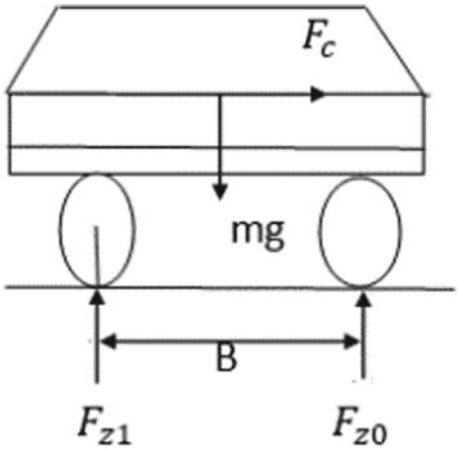


Fig. 5 Analysis of forced vehicle turning

And get

$$u \leq \sqrt{gR\varphi_y} \tag{2}$$

According to formula 2, as the turning radius and lateral adhesion coefficient increase, the speed of the vehicle also changes. This article considers that the vehicle's anti-sliding capability can be represented by a critical speed indicator without slipping. When the road has a lateral slope angle of α, when the vehicle makes a turning motion, the condition of the side slope of the turning inclination does not occur is:

$$u \leq \sqrt{\frac{Rg\left(\varphi_y \pm \tan\alpha\right)}{1 - \varphi_y \tan\alpha}} \tag{3}$$

In Eq. (3), + indicates that the road faces the inside of the curve, and—indicates that the road is inclined toward the outside of the curve. With the road parameters unchanged, with reference to the forward direction of the vehicle, the anti-sliding stability of the inner slope of the road curve is better than that of the outer side.

3.2 Anti-rollover Stability Model

This paper studies the quasi-static rollover of a rigid vehicle, as shown in Fig. 6, when the vehicle turns, due to the different points of lateral force, the generated torque causes the vehicle to roll to the outside of the curve [11].

Analysis of the turn process, assuming an ideal state: when the vehicle is turning, it will move in a nearly uniform circular motion, and the vehicle will not destabilize.

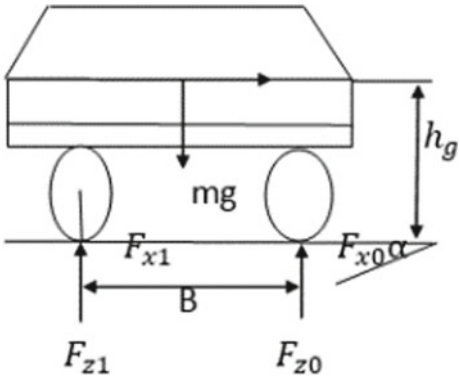


Fig. 6 The force acting on a vehicle to roll it over

The lateral ramp angle should also be considered on the freeway. When a vehicle makes a turning motion, it is available according to the analysis of force balance:

$$ma_y h_g - mgh_g \sin\alpha + F_{z1} B - \frac{1}{2} mgB\cos\alpha = 0 \tag{4}$$

In the formula: α is the lateral ramp angle of the road, m is vehicle quality, h_g is the height of the center of mass of the vehicle, B is track distance, a_y is lateral acceleration when the vehicle is steadily steered, F_{X0}, F_{X1} are lateral forces on the inner and outer wheels of vehicles, F_{Z0}, F_{Z1} are the normal reaction force of the ground on the inner and outer wheels of the vehicle, R is turning radius, V is driving speed. Organize Formula (4) to get:

$$\frac{a_y}{g} = \left(\frac{1}{2}\cos\alpha - \frac{F_{z1}}{mg}\right)\frac{B}{h_g} + \sin\alpha \tag{5}$$

When there is no lateral acceleration and horizontal road surface, when the load on the inner wheel is equal to half of the total weight of the vehicle, it should satisfy:

$$\varphi = \frac{a_y}{g} \tag{6}$$

According to formula 6, the size of F_{Z1} can be used to balance the rollover moment generated when turning. Considering the most dangerous conditions—when turning at extreme speeds, the wheels on the outside of the vehicle are subject to all weights, and the wheels on the inside of the vehicle cannot balance the roll, and rollover in the event of instability occurs. See the formula below.

$$\frac{a_y}{g} = \frac{B}{2h_g}\cos\alpha + \sin\alpha \tag{7}$$

The lateral slope of the road is very small and can be approximated as $\sin\alpha = \alpha$, $\cos\alpha = 1$. Finishing formula (7):

$$\frac{a_y}{g} = \frac{B}{2h_g} + \alpha \tag{8}$$

Take $\alpha = 0$ in the actual situation. The rollover threshold at this time is $\frac{B}{2h_g}$ to evaluate the rollover resistance of the vehicle. The threshold increases significantly with increasing wheelbase and decreasing centroid height. This paper considers the available vehicle speed to characterize the rollover condition of the vehicle, and uses the critical safety speed of the vehicle to avoid the rollover accident of the vehicle. When a vehicle makes a steady circular motion on a horizontal road surface, the vehicle does not roll over:

$$\mathrm{u} \leq \sqrt{\frac{BRg}{h_g}} \tag{9}$$

The risk of rollover of the vehicle is much greater than that of the side slide. To ensure that the side slide occurs before the rollover, the vehicle can be fully protected against rollover accidents. In summary, the analysis of the vehicle's anti-skid and anti-rollover movements meets the following conditions:

$$\sqrt{Rg\varphi_y} \leq \sqrt{\frac{BRg}{2h_g}} \tag{10}$$

Organize Formula (10) to get:

$$\varphi_y < \frac{B}{2h_g} \tag{11}$$

It can be seen that the greater the wheelbase B, the smaller the center-of-mass height h_g of the vehicle, the greater the lateral adhesion coefficient of the road surface, and the more secure the lateral stability of the vehicle.

4 Rainy Day Highway Safety Simulation

4.1 Establishment of ADAMS/Car Model Under Rainfall Conditions

This paper adopts software ADAMS/Car to establish road model, vehicle dynamics model and vehicle-road coupling model, using this as a means to simulate the driving

conditions of the vehicle under dry asphalt and rainy road conditions, analyze the response of vehicles driving in a rainy environment according to the dynamics and kinematic response of the vehicle.

4.2 Design of Simulation Program for Driving Safety on Expressway at Rainy Days Based on ADAMS/Car Analysis

The simulation scheme design is the basis of the simulation experiment module. It selects the required simulation for the user and selects the corresponding simulation scheme for different simulation requirements. To study the critical safety speed of car turning and overtaking stability under different rainfall intensity levels. In this paper, the maximum rainfall intensity under different rainfall weather is calculated and the simulation parameters are designed for experimentation. The water film thicknesses during light rain, moderate rain, heavy rain, and heavy rain are 0.63, 1.547, 2.647, 5.001 mm respectively.

Experiments on steering and double shifting of vehicle under different rainfall intensity, observing the output of lateral acceleration a_y and yaw rate w_r of the vehicle while the vehicle is running. If the data is within the threshold, it can be determined that the vehicle is in a safe state, otherwise the vehicle is in an unsafe state.

Looking at the design rules of the Chinese highway route, we can see that the minimum radius of round curves at all levels of highways is shown in Table 1.

Critical speed has the meaning of limit value, this paper selects the limit value of 400 m for the minimum radius of the circular curve of the highway design speed 100 km/h as the research object, study the turning conditions of the vehicle, analyze the critical condition of the vehicle and get the limit of different road conditions. Simulation speed is 60–120 km/h, step length is 10 km/h, it indicates that the vehicle has skid or rollover until the output response of the vehicle does not meet the safety assessment indicator. The specific combination is shown in Table 2.

Table 1 The minimum radius of round curves of roads at all levels

Design speed (km/h)	120	100	80	60	40	30
Limit value (m)	650	400	250	125	60	30
General value (m)	1000	700	400	200	100	65

Table 2 Turning simulation program combination

	60 km/h	70 km/h	80 km/h	90 km/h	100 km/h	110 km/h	120 km/h
400 m							

4.3 Vehicle Safety Driving Judgment Parameter at Rainy Day

The yaw rate, lateral acceleration, and lateral displacement speed were selected as criteria.

(1) Yaw rate W_r, slip angle α

Refer to previous studies and this article on vehicle rollover, anti-slip model construction, and vehicle stability determination parameters [12–14], the yaw rate of the vehicle during driving will not exceed 0.15 rad/s, the side slip angle is less than 0.10 rad, within such a safety threshold, the vehicle is in a steady state. When the slip angle fluctuates in a very small range, the yaw rate plays a decisive role. This article suggests choosing $W_r \leq 0.15\,\text{rad/s}, \alpha \leq 0.10\,\text{rad/s}$.

(2) Lateral acceleration α_y

For cars, there is $\frac{B}{h_g} = 1$ at full load, Combining with the analysis of driving stability of the vehicle, the critical condition of preventing the occurrence of sideslip and rollover of the vehicle is satisfied: $a_y = \varphi_y < \frac{Bg}{2h_g} < 0.5g$.

(3) Lateral displacement speed v_y

When the vehicle on the road is destabilized, observing the lateral displacement curve, it is found that the curve is not continuous, it is fragmented or incomplete, or it has a large change in the adjacent time period.

5 Analysis of Simulation Results of Safe Driving on Expressway in Rainy Day

5.1 Analysis of Simulation Results

Simulated analysis of fixed turning conditions and double shifts, limited by the space relationship, only the results of simulation analysis under fixed turning conditions.

Under a constant turning condition with a radius of 400 m, simulations are performed according to the simulation scheme. It can be seen from the analysis of the output curve of the evaluation indicators, the state of the vehicle is divided into two states of instability and instability, with regard to the analysis of these two states, this article selects one case to illustrate.

(1) The vehicle did not lose stability

The analysis of the vehicle's unsteady state is carried out, and the vehicle state is analyzed at a speed of 104 km/h in heavy rain conditions. The simulation parameters are shown in Table 3.

The lateral acceleration, lateral displacement, and yaw rate of the vehicle without stabilizing state are shown in Fig. 7a–c.

Table 3 Simulation parameters

Simulation parameters	Simulation time (s)	Speed (km/h)	Turning radius (m)	Rainfall rating	Adhesion coefficient
Value	60	104	400	Heavy rain	0.236

From the output response analysis in Fig. 7, we can see that, when the speed is 104 km/h in heavy rain, the vehicle's lateral acceleration value and yaw rate value when the vehicle is not in a state of stability are all satisfying the threshold of the vehicle's stable driving evaluation index. From the smooth and complete analysis of the lateral displacement curve of the vehicle, it can be seen that the vehicle is in a safe state at the moment.

(2) Vehicle instability

Analyze the state of vehicle instability, select the vehicle speed at a speed of 110 km/h for heavy vehicle conditions. The simulation parameters are shown in Table 4.

The lateral acceleration, lateral displacement, and yaw rate of the vehicle in a destabilizing state are shown in Fig. 8a–c.

From the output response analysis, when the vehicle speed is 110 km/h in heavy rain, the vehicle's lateral acceleration, yaw rate, and lateral displacement curve are incomplete and drastic changes when the vehicle is in a state of instability, and the vehicle will be uncontrolled without stability.

In the case of light rain, moderate rain, heavy rain, and heavy rain, simulate vehicle speeds from 60 to 120 km/h, and take 10 km/h steps to find the critical safety speed value of the vehicle through the output response. Due to space limitations, only the different vehicle speeds under heavy rain conditions are used for vehicle status analysis.

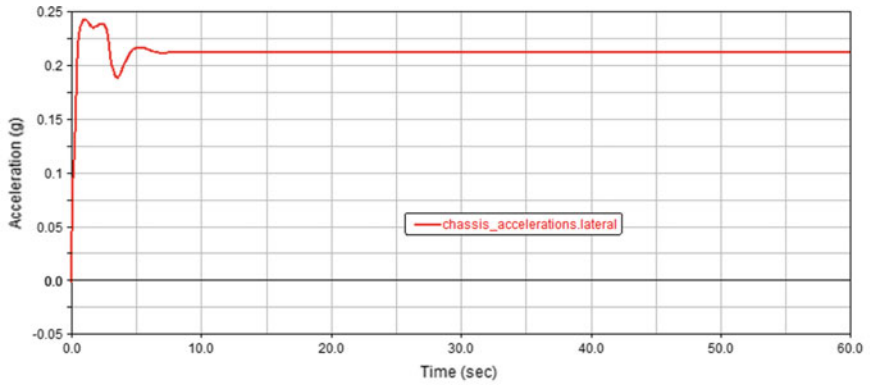
Analyze the safety status of the vehicle and select vehicle speeds under different conditions of heavy rain. The simulation parameters are shown in Table 5.

The lateral acceleration, lateral displacement, and yaw rate of the vehicle under heavy rain conditions are shown in Fig. 9a–c.

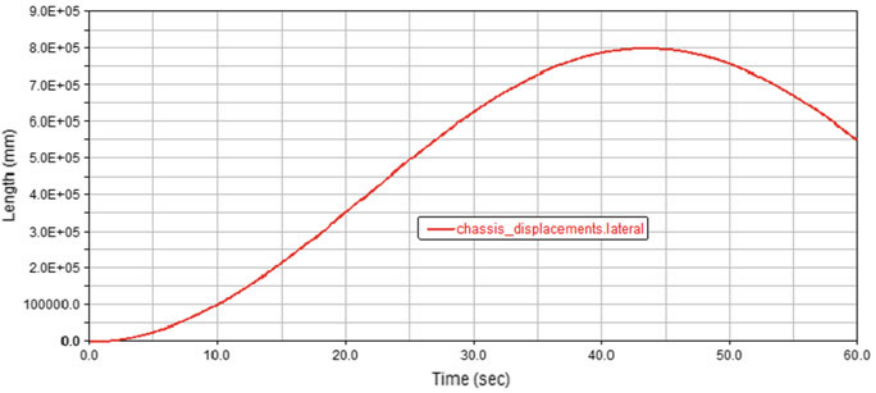
It can be seen from the output response analysis in Fig. 9, when the vehicle speed is 110 km/h in heavy rain, the vehicle's lateral acceleration, yaw rate, and lateral displacement curve are incomplete and produce drastic changes, the analysis shows that the vehicle is in an unsafe state at 5.0 s and the vehicle will be uncontrolled without stability. Through the above analysis, it can be seen that vehicle instability occurs at 100–110 km/h in heavy rain.

Under severe rain conditions, a simulation speed was calculated at a speed of 100–110 km/h with a step length of 1 km/h, so as to obtain a critical safety speed value, as shown in Fig. 10.

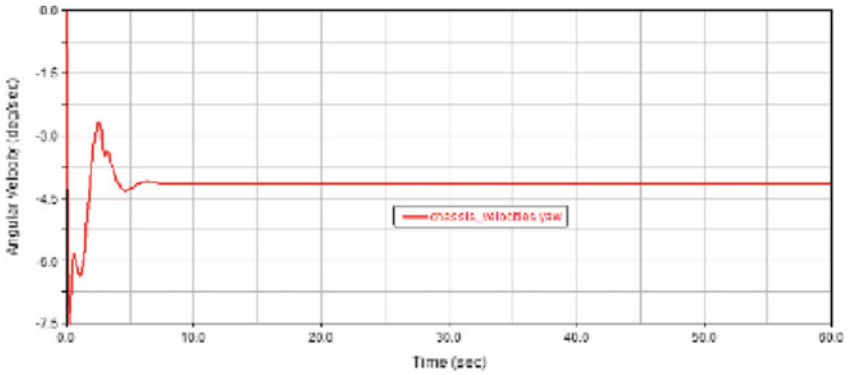
It can be seen from the analysis of Fig. 10, the red curve is the driving output response of the vehicle when the speed is 104 km/h, the blue curve is the vehicle output response with a vehicle speed of 105 km/h, The yaw rate sharply changes to 100 deg/s, greater than 0.15 rad/s; Lateral acceleration is changed from +0.2 to − 0.2 g, the movement trajectory is offset, and the vehicle is in a state of instability.



（a）Lateral acceleration



（b）Lateral displacement



（c）Yaw rate

Fig. 7 The vehicle does not turn stably output response

Table 4 Simulation parameters

Simulation parameters	Simulation time (s)	Speed (km/h)	Turning radius (m)	Rainfall rating	Adhesion coefficient
Value	60	110	400	Heavy rain	0.259

Therefore, the vehicle's critical safety speed is 104 km/h and the adhesion coefficient is 0.236.

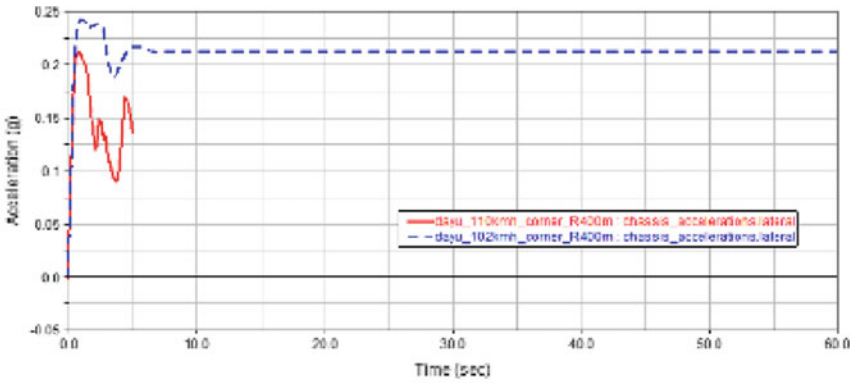
In order to obtain the critical safety speed values for vehicles with different rainfall intensities, the results are shown in Table 6.

5.2 Simulation Results Verify Reliability

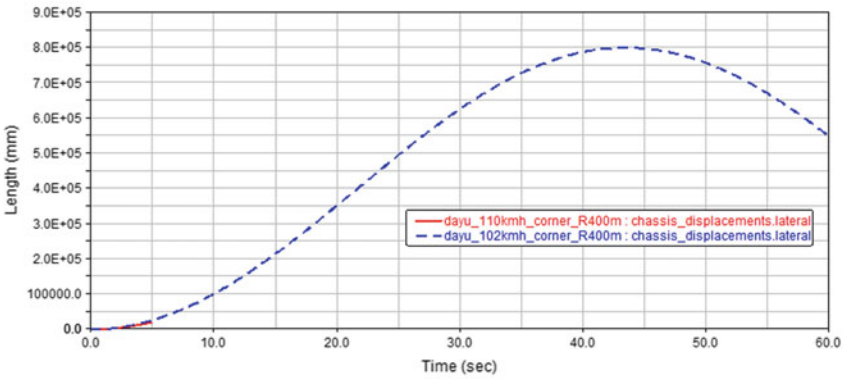
Substituting critical safety speed into braking distance formula, the braking distance can be obtained as shown in Table 7. The analysis shows that under light rain, moderate rain, heavy rain, and heavy rain conditions, the friction coefficient of the road determines the driving safety of the vehicle, the braking distance of the critical safety speed is within the visibility range, and the critical speed meets the stability of the vehicle. Therefore, it is reliable to determine the safe speed value by integrating the two indexes of road friction coefficient and visibility.

6 Conclusions

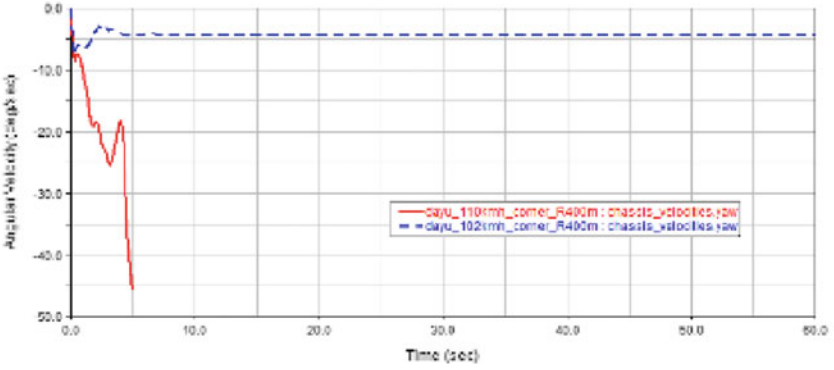
This paper studies the impact of rainy weather on the driving stability of the vehicle. The dynamic analysis of the vehicle establishes the vehicle's anti-slip and anti-rollover model, and concludes that the vehicle does not experience a side-slip condition when the vehicle is turning, and the vehicle does not roll over. The paper also analyzes the phenomenon of hydroplaning in rainy days, and discusses the relationship between water film thickness, adhesion coefficient and vehicle speed. Based on vehicle stability analysis, vehicle safety driving evaluation parameters are obtained, and ADAMS/Car simulation is used to determine the safety characteristics of vehicles in rainy days. Analyze the maximum safe driving speed on Expressway in rainy days. In view of the fact that the predecessors have less research on the speed limit of vehicle stability in the rainy weather environment, this study supplements this and also provides theoretical support for the vehicle's reasonable speed limit under adverse weather conditions.



（a）Lateral acceleration



（b）Lateral displacement

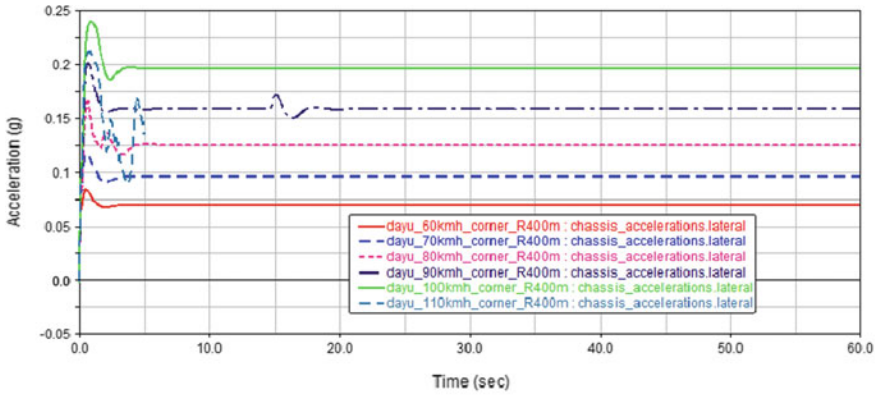


（c）Yaw rate

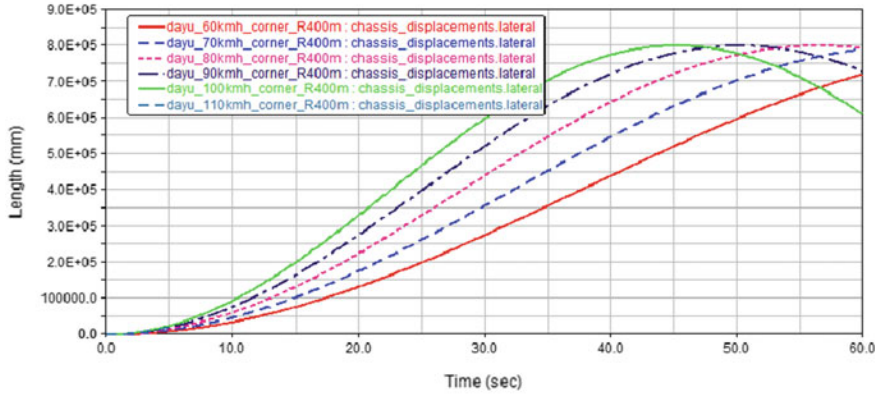
Fig. 8 Vehicle's fixed turn instability output response

Table 5 Simulation parameters

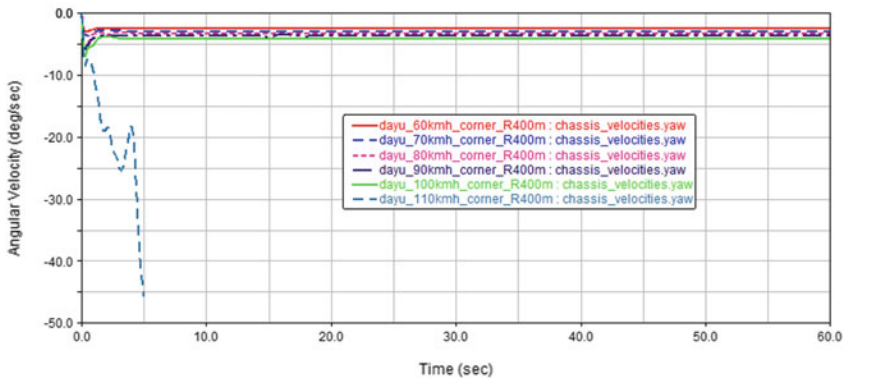
Simulation parameters	Simulation time (s)	Speed (km/h)							Rainfall rating
		60	70	80	90	100	110	120	
Value	60	Adhesion coefficient							Heavy rain
		0.4870	0.4300	0.3730	0.3160	0.2590	0.2020	0.1450	



（a）Lateral acceleration

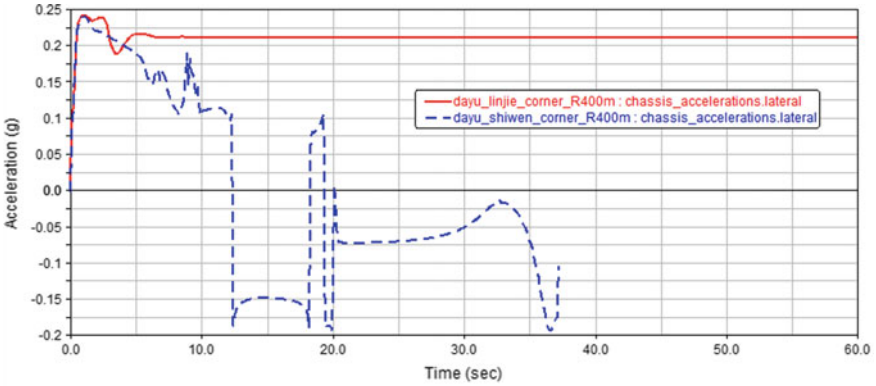


（b）Lateral displacement

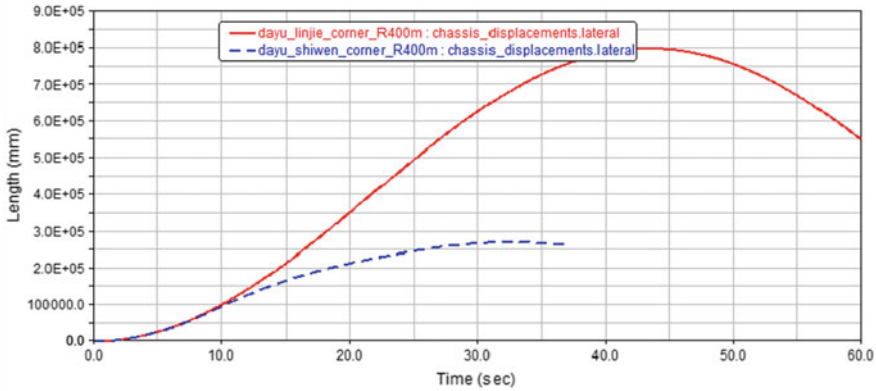


（c） Yaw rate

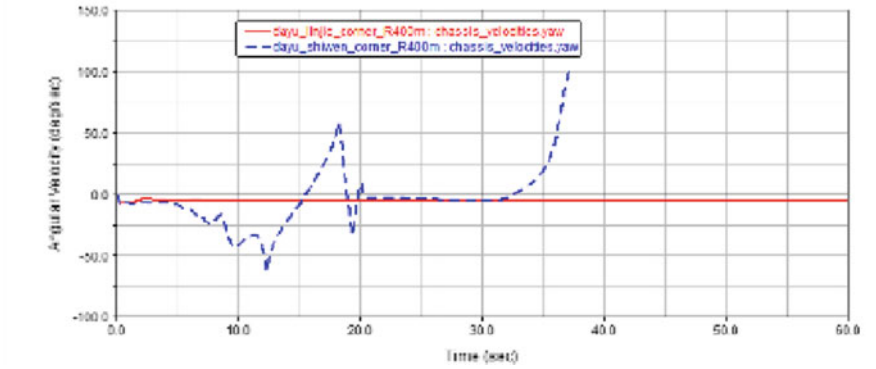
Fig. 9 Heavy rain turns output response



（a）Lateral acceleration



（b）Lateral displacement



（c） Yaw rate

Fig. 10 Turning critical safety speed output response in heavy rain

Table 6 vehicle safety critical speed value

Rainfall intensity	Road adhesion coefficient	Critical safety speed (km/h)
Light rain	0.280	112
Moderate rain	0.256	109
Heavy rain	0.236	104
Rainstorm	0.263	81

Table 7 The maximum distance of driving and braking on expressway in rainy day

Weather level	Critical speed (km/h)	Maximum braking distance (m)	Visibility may decrease (m)
Light rain	112	165	500
Moderate rain	109	168	250
Heavy rain	104	150	80–150
Rainstorm	81	92	50–100

References

1. Andrey J, Yagar S (1993) A temporal analysis of rain-related crash risk. Accid Anal Prev 25(4):465–472
2. Cardiology WGOC (2001) Recommendations for exercise training in chronic heart failure patients. Eur Heart J 22(2):125–135
3. Smith K (2008) How Seasonal and weather conditions influence road accidents in glasgow. Scott Geogr J 98(2):103–114
4. Hooper E, Chapman L, Quinn A (2014) The impact of precipitation on speed–flow relationships along a UK Motorway Corridor. Theor Appl Climatol 117(1–2): 303–316
5. Yin T, Jia X (2009) Study on safe driving speed of rainy day on freeway based on parking distance. Northern Commun 1:150–152
6. Ji T, Huang X, Liu Q et al (2004) Road surface water film thickness prediction model. J Traffic Transp Eng 4(3):1–3
7. Li H, Chen G, Xia Z (2009) Reasonable value of maximum speed limit in bad weather conditions on the highway. J Shijiazhuang Railway Institute (Natural Science) 22(3): 78–81
8. Huang W (2003) Vehicle skidding and disposal. Labour Prot 11, 76–77
9. Li M (2011) Research on algorithm for estimation of side tilt angle and road adhesion coefficient of wire-controlled vehicle tires. Jilin University
10. Jin Z, Zhang H, Ma C (2012) Dynamic stability based vehicle rollover warning. Chin J Mech Eng 48(14):128–133
11. Pan K (2009) Research on the driving state of the second and third grade highways in mountain areas. Chang'an University
12. Horne WB (1968) Tire hydroplaning and its effects on tire traction. High Res Rec
13. Mounce JM, Bartoskewitz RT (1993) Hydroplaning and roadway tort liability
14. Ewing R, Cervero R (2001) Travel and the built environment; a synthesis. Transp Res Rec 265–294

Optimization Method of Comb-Shaped Speed Reduction Marking Spacing Based on Variable Space-Time Frequencies

Liangjie Xu, Zhijun Wang, Ruonan Zhou and Hua Fan

Abstract In view of the practical matters that drivers often have traffic accidents on the freeway curves because of speeding, it is essential to improve the traffic security level and reduce the traffic accident rate of highway. And the key aspects are to keep a safe running speed or let the speed down by setting up reduction measures like speed reduction marking. This paper presents a more scientific design scheme of speed reduction markings. Taking comb-shaped speed reduction marking for example, it's spacing were optimized for further research. First, this paper gives explanations on mechanism of the space-time frequency influence on the drivers' perception of vehicle speed from the angle of body sensing (mainly visual sense). Second, a mathematical model to describe the relationship between spatial frequency and spacing of speed reduction marking was established. After defining comfortable acceleration, maximum safe speed, etc. aiming at different freeway curves, the model can obtain specific value of speed reduction marking's spacing. At last, to explore the practical effects on drivers' speed perceptions, this study simulates highway scenes by 3Dmax. Those road surfaces are patterned with speed reduction markings set in different space-time frequencies. And then evaluates the deceleration effects of those speed reduction markings by comparative analyzing experimentalists' reactions recorded. The experimental results shows that the deceleration effect can be improved effectively by decreasing the spacing of speed reduction markings, namely increasing the frequency. In addition, variable space-time frequencies give drivers illusions that the running speed keeps changing. In general, speed reduction markings set in variable space-time frequencies have better effects on speed reduction.

Keywords Traffic safety · Freeway · Speed reduction marking · Speed perception · Space-time frequency

L. Xu · Z. Wang · R. Zhou (✉) · H. Fan
School of Transportation, Wuhan University of Technology, Wuhan 430063, Hubei, China
e-mail: 18696172619@163.com

Z. Wang
Planning Department, Jilin Province Highway Construction Bureau,
Changchun 130033, Jilin, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_26

1 Introduction

The freeway curves are accident prone locations. In 2014, the amount of traffic accidents in China was 70,065, which reached 35.60% of the total number of accidents [1]. It takes serious damage on life and wealth of people, causing terrible consequences. Speeding is the major factor in this issue. In order to decrease the accident rate on highway effectively, we need to control the speed within a safe speed range.

In recent years, a growing body of research has concluded speed perception at home and abroad. One sampling survey found that about half of the drivers were running faster than their perceived speed [2]. Speed perception is fundamental for drivers' behavior to accelerate or decelerate. In line with this, edge rate is one important reference substance or stimulant to sense the velocity, which was put forward in 2003 [3]. It refers to the number of edges or gaps that pass through the observer's visual field within a unit time, a kind of information about time frequency. Domestic and foreign scholars express speed as the ratio of time frequency and spatial frequency. The time frequency refers to the number of times that the stimulus appears per unit time. The spatial frequency is the number of intervals per unit length, which is the reciprocal of the interval length [4]. When the observer's movement speed is certain, if the spacing of the texture is narrowed, the observer's perception speed will increase [5]. Besides the trees and guardrails on both sides of the road, Speed reduction markings are widely used to influence the speed perception of drivers and induce the driver to slow down consciously. Most of the research around the speed reduction marking and edge rate are field testing to detect the deceleration effect of some constant frequencies. In this way, we can only get drivers' first reaction to test markings, ignoring the ability of self-adjusting and speed adaptability of human body. When the drivers get familiar with the road condition, the deceleration effect will only get worse. But the gradual narrowing spacing are assumed to give drivers illusion that the running speed is increasing gradually and offset the adaptability of the speed.

In order to reduce the effects of speed adaptability, based on the space-time frequency, this paper establishes mathematical model between spatial frequency and speed reduction marking spacing. Taking comb-shaped speed reduction marking for example, which has the best deceleration effect [6], compare the real deceleration effect of different spatial frequencies by simulation experiment.

2 Related Definition and Influence Mechanism

Time frequency refers to the number of times that the stimulus appears per unit time. Spatial frequency is the number of intervals per unit length. During the driving process, speed reduction marking is the stimulus. By changing the spatial frequency of the marking spacing, the time frequency that driver subjectively perceives is accordingly changed while moving. Thus the driver's perception of speed is affected.

According to the definition of time frequency and space frequency, the physical velocity is the ratio of time frequency to spatial frequency:

$$V = \frac{f_t}{f_s} \tag{1}$$

where

V refers to physical velocity;
f_t refers to time frequency;
f_s refers to spatial frequency.

Shen [7] obtained the mathematical model of perceptual speed with time and spatial frequency by fitting the measured data:

$$V_p = v^{\alpha} \cdot f_t^{1-\alpha} = \frac{v}{f_s^{\alpha-1}} \tag{2}$$

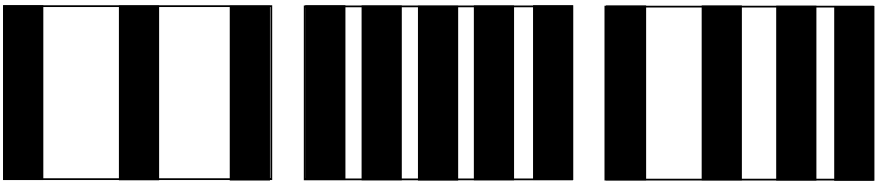
where

V_p refers to time frequency perceived speed;
α refers to time frequency the parameter obtained by fitting.

Different frequencies give drivers different information about speed. To reduce the influences of speed adaptability and enhance the markings' reduction effects, the spacing between deceleration markings need to be narrowed, namely, the spatial frequencies should be go up gradually, as shown in Fig. 1.

There is a clear effect of driving environment on perceived speed. The mechanism that variable spatial frequencies reduce speed is shown in Fig. 2.

When the driving environment changes, the surrounding traffic information changes with it. And they change the driver's perception speed, so as to induce their driving behavior to slow down. During the moving process, variable spatial frequencies can cause the change of driver's driving behavior even more effective. While the frequency gets bigger, the driver can clearly sense that it takes less and less time to see the next marking, which means the speed that fringes pass through drivers' field of vision gets faster. And furthermore, the monotonous highway environment and wide visual field of driving make it impossible for the drivers to tell



(a) Low frequency. (b) High frequency. (c) Variable frequency.

Fig. 1 Different spatial frequencies

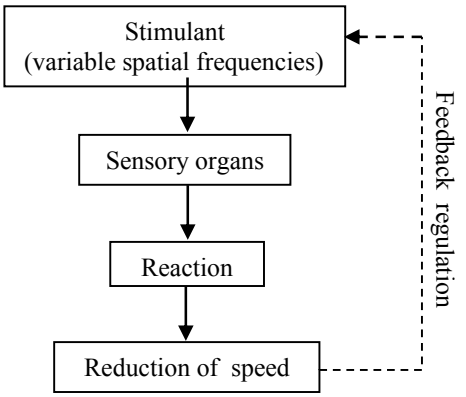


Fig. 2 Speed control mechanism

the root cause of this change. Above all, the specific reasons why speed reduction markings can reduce speed are as follows:

(1) Increasing time frequency. The narrowing marking spacing increases the number of speed reduction marking passing through during the same time. So that the drivers overestimate the speed and then reduce the speed;
(2) Overcoming speed adaptability. The speed reduction markings are regarded as stimulant. However the impacts of constant stimulus intensity on sensory body organs are going to diminish or even disappear. The same as speed perception. Speed adaptive makes drivers underestimate speed. The changing spacing give drivers Continuous stimulus, eliminating the driver's speed adaptability.

3 Design Methods

3.1 Design Concept

According to the concept of spatial frequency, If the texture element spacing is known, markings with different spatial frequencies that appear in the drivers' view can tell drivers speed. When the spacing between the texture elements narrow down, the speed is perceived to be higher [8]. When drivers overestimate the speed, their bodies' self-protection mechanisms will perceive a risk. Then they will uniformly decelerate with a relatively comfortable and safe deceleration. According to American Association of State Highway and Transportation Officials (AASHTO)' determination of the driver's comfortable deceleration, set comfortable deceleration to -3.0 m/s^2 [9]. So, in this paper, the basic assumption of the design is that the drivers will decelerate with the comfort deceleration (-3.0 m/s^2) after the they find they're speeding and feel unsaved.

Here is the design idea:

① Investigate the maximum speed among the curve accident on highway, take it as initial speed of entering the test deceleration road.
② Investigate the terrain conditions, calculate the maximum safe speed according to the curve radius, the high elevation and the transverse force coefficient of the highway, take it as desired speed while passing through the test deceleration road.
③ The spatial frequencies of speed reduction markings' spacing decrease uniformly along the driving direction.

3.2 Design Model

Assume that vehicles get into the sections with initial speed v_0, length of the section is l, driven out of the road at the desired speed v_t, then get into that curves which radius is R, ultra high is i_h, and the transverse force coefficient is μ, time is t, and the width of the marking is 0.15 m, and the length is 0.5 m, then:

$$v_t = \sqrt{127R(i_h \pm \mu)} \tag{3}$$

$$l = \frac{v_t^2 - v_0^2}{2a} \tag{4}$$

where

v_0 refers to 85% speed of 200 m before entering the highway curve according to historical data.
v_t refers to maximum safe speed before entering the highway curve.
a refers to comfort deceleration (-3.0 m/s^2).

According to Formulas (3) and (4), the expected speed and the length of the road section laying with the speed reduction markings are known.

Experimental results show that under the speed of 120, 140, 160 km/h, edge rate between 4 and 16 Hz, all of the experimenter overestimated the speed, and with the increasing of the overestimation effect of edge rate density, the difference in the driver's speed perception is small. And with the increase of edge rate, the overestimation effect gradually increases. However, when the edge rate density is 8–16 Hz, it's not obvious that the change of driver's perceived speed [10]. According to the experimental conclusion, pave comb-shaped speed reduction marking with spatial frequency between 8 and 16 Hz (the speed is 120 km/h) in the road section, as shown in Fig. 3.

Suppose that n markings is paved, namely there are n − 1 spacing needing to be reasonably designed, set x_i as the value of the ith spacing, then:

$$l = 0.5n + \frac{(n-2)(x_1 + x_{n-1})}{2} \tag{5}$$

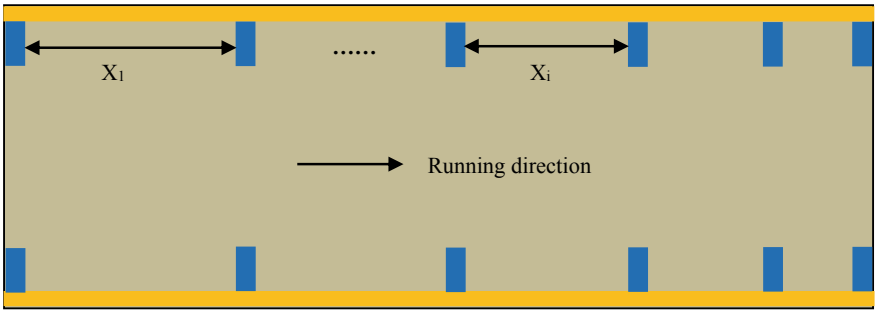


Fig. 3 Comb-shaped speed reduction marking based on variable space frequency

$$d = \frac{x_n - x_1}{(n-1)} \tag{6}$$

$$x_i = x_1 + (i-1)d \tag{7}$$

where

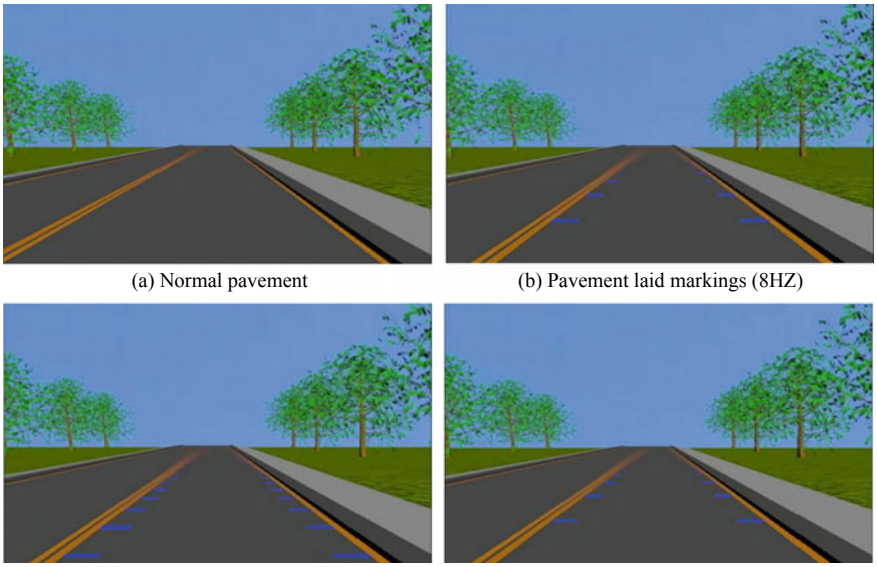
d refers to the difference between adjacent spacing.
x_i refers to the corresponding spacing under the speed of 120 km/h and the spatial frequency of 8 Hz, 4.02 m.
x_n refers to the corresponding spacing under the speed of 120 km/h and the spatial frequency of 8 Hz, 1.83 m.

According to the Formula (5), it can be calculated that the number of markings paved on the road. According to the Formula (6), the difference of two adjacent markings spacing can be calculated. According to the Formula (7), we can calculate the each specific value.

4 Simulation Experiment

4.1 Apparatus

Use computer graphics software (3Dmax 2016) to make virtual traffic environment for video production in the experiment process. And TOSHIBA projector was used to play video.



(a) Normal pavement (b) Pavement laid markings (8HZ)

(c) Pavement laid markings (16HZ) (d) Pavement laid markings (8~16HZ)

Fig. 4 Different road model

4.2 Observers

20 experimenters were selected (all with driving experience, naked eye vision or corrected visual acuity not less than 1.0). Record the their reactions (velocity and velocity changes) of the observer during the experiment.

4.3 Scene Model

In order to simulate the actual road conditions, the road model is two-way two-lane road, the lane-width is 3.75 m, and the lane-line-width is 0.15 m. The total length of the section is 100 m. And the markings are 0.6 m in length, 0.3 m in width, as shown in Fig. 4.

4.4 Procedure

3D max software is used to make virtual video that vehicles are running with the speed of 80, 100 and 120 km/h on the highway, which surfaces are paved with different speed reduction markings designed in 8, 16, and 8–16 Hz variable spatial frequencies.

Table 1 Spacing under specific speed and spatial frequency condition

Speed (km/h)	Spatial frequency (Hz)	Spacing frequency (m)
80	8	1.24
	16	2.63
100	8	1.59
	16	3.32
120	8	1.83
	16	4.02

In addition, control groups (normal road without changing) are given. Videos that driving on the normal road and designed road are played simultaneously. Take speed in control groups as standard stimulus. Under the same speed condition, there are 9 groups of comparison schemes. Let the observers sense the speed and variety of speed by watching the video. "Slower, the same, faster" are provided to describe speed. And "+" is used to express degree. "Slow down, keep constant and speed up" are provided to describe variety of speed. Record their feelings and perceptions.

According to spatial frequency and speed, the spacing are known. Table 1 provides each spacing under specific speed and spatial frequency condition.

4.5 Data Analysis of the Experiment

Experimental data were collected and processed, as shown on Table 2.

From data statistics and analysis, we can come to the following conclusions:

Table 2 Three group's experimental data

Experimental condition		Speed			Variety of speed		
Speed (km/h)	Frequency (Hz)	Faster +	Faster + +	Faster + ++	Slow down	Keep constant	Speed up
80	8	18	2	0	0	20	0
	16	0	5	15	0	20	0
	8–16	0	10	10	0	0	20
100	8	14	6	0	0	20	0
	16	0	3	17	0	20	0
	8–16	0	5	15	0	0	20
120	8	12	8	0	0	20	0
	16	0	2	18	0	20	0
	8–16	0	4	16	0	0	20

(1) With the improvement of the experimental speed, the speed perception of the driver improved to varying degrees.
(2) With the increase of spatial frequency, the overestimation effect gradually increases.
(3) Compared with the constant spatial frequency of 8 and 16 Hz, the variable spatial frequency of 8–16 Hz can give the driver an illusion of speed's variety.

5 Conclusion

This study regards straight-line before entering highway curves as research object, does optimal designs of comb-shaped speed reduction marking's spacing, considers the influence of time frequency and space frequency on driver's speed perception, puts up with the principle of marking spacing and the design method, establishes the mathematical model spatial frequency and speed, obtains the following conclusions through simulation experiment:

(1) The driver's perceived speed can be improved effectively by laying markings with variable spatial frequency.
(2) markings with variable spatial frequency can not only increase the perceived speed but also overcome the speed-adaptability.
(3) It is necessary to take road alignment into consideration when designing the speed reduction marking.

References

1. Han J (2014) The concept and method of designing highway in high altitude cold region and application. Chang 'an University
2. Haglund M, Aberg L (2000) Speed choice in relation to speed limit and influences from other drivers. Transp Res Part F: Traffic Psychol Behav 3(1):39–51
3. Wickens CD, Hollands JG (2003) Engineering psychology & human performance. East China Normal University Press, 193–196
4. Shen H, Shimodaira M, Ohashi G (2005) Speed-turned mechanism and speed perception in human vision. Syst Comput Jpn 36(13):1718–1727
5. Andersen GJ, Cisneros J, Atchley P, Saidpour A (1999) Speed, size, and edge-rate information for the detection of collision events. J Exp Psychol. Hum Percept Perform 25(1)
6. Zhu X, Zhao W, Liu H, Liu X, Zhang G (2012) A comparative study on the effect of simulated highway visual illusion deceleration 01
7. Shen H M, Shimodaira, Ohashi G (2005) Speed-turned mechanism and speed perception in human vision. Syst Comput Jpn 36(13):1718–1727
8. Richard RA, McGee HW, Farmer CM (2000) Influence of experimental pavement markings on urban freeway exist-ramp traffic speeds. Transp Res Rec 1705, paper No. 00-3272

9. AASHTO (2001) American Association of State Highway and Transportation Officials (AASHTO). A policy on geometric design of highways and streets
10. Liu B (2008) Study on vehicle speed control and lane-keeping mechanism based on driver visual perception. Wuhan University of Technology

Summary of Research on Exit Control of Urban Expressway

Yan Xing, Shuai Bian, Wei-dong Liu, Xing-quan Guan and Yang Liu

Abstract With the rapid development of urban expressway, the ramp congestion of expressway is frequent, and the ramp control of expressway exit has been widely concerned. There have been more research and application on the control of expressway exit in China, and the control mode is mainly focused on the adjustment of the adjacent auxiliary road intersection on the ramp. In order to continue to find ways to solve the problem of congestion on exportation of expressways, and provide reference for the research of expressway exit control, the existing control methods are summarized. Although most of the exportation control is adopted to adjust the signal of the auxiliary road at present, the object of the control is not the same. Therefore, the object different according to the control means can be divided into three kinds: the priority of the ramp, the dynamic adjustment of the auxiliary road and the cooperative control of the multi system. This article will summarize and analyze the existing research results, and put forward some views on the future research and development.

Keywords Express road · Ramp · Auxiliary road · Coordinated control · Cooperative control

1 Introduction

The main reason for the congestion at the ramp exit of urban expressway is that the urban expressway is mostly built in the urban road directly, and the surrounding road facilities and functions have been fixed stationary, which leads to the inability of the

Y. Xing · S. Bian · W. Liu (✉) · X. Guan · Y. Liu
School of Transportation Engineering,
Shenyang Jianzhu University, Shenyang 110168, Liaoning, China
e-mail: 44667276@qq.com

Y. Xing
e-mail: 359063088@qq.com

Y. Xing
School of Automotive Engineering, Jilin University, Changchun 130025, Jilin, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_27

entrance and exit of the expressway that can not completely avoid the other road intersection, causing the vehicles to operate at the entrances and exits of the ramps are disturbed by other road vehicles.

At present, the main way to control the exit of urban expressways is to exit ramps and auxiliary road regulation which near the exit ramps [1]. The off ramp closure was first implemented in the management of expressways. Due to the disadvantages of practical application, it has been used less. Most of the control methods used now are adjusting the adjacent auxiliary roads of the expressway exit ramp, and signal period is changed to the dynamic model which is constantly adjusted with the change of the traffic flow state. Thus the change of the traffic flow can be felt in time, and the traffic flow congestion problem of the ramp is alleviated. Many researches have also been made on the latter, and the research contents are quite rich. In this paper, the different control modes of expressway exit are divided into three categories. The three main types of export ramp control are summarized as follows: priority of traffic flow on exit ramp, dynamic adjustment of auxiliary road, and coordinated control of multiple systems.

2 Control Thought

The essence of the expressway exit congestion problem is the supersaturated traffic flow at the exit ramp. In order to dredge the vehicle as soon as possible while not obstructing the traffic of other road vehicles, the commonly used management method is to change the signal control scheme of the auxiliary road, by limiting the signal timing of the auxiliary road flow through or adjusting the auxiliary road. To speed up the release of traffic flow on the expressway.

When the control method of expressway exit is chosen, it is necessary to consider whether the state of the road vehicle after the control is better than the state before the control, so that the expressway and other roads can reasonably distribute the traffic flow, make the road network achieve the best benefit, and analyze whether there is enough space for queued vehicles to meet the demand for parking in queuing vehicles Then ensure that the length of the queuing vehicles will not be too long and will be traced back to the expressway. It should be considered that after the use of control means, there is any significant improvement in the operating status of the expressway, such as the length of the queue, the time of queuing and other, and the other corresponding indicators should also have a good trend of development.

In order to release the traffic jam on the ramp quickly, the main control means in China can be divided into three categories: In the first category, priority is given to traffic on the exit ramp: when traffic passes through the auxiliary road intersection in the exit direction of the expressway, the traffic light will preferentially release the vehicles in that direction, preventing the vehicle.

3 Control Plan

Many research achievements can be summed up in the following aspects: priority of traffic flow in the export direction of the ramp: higher degree of application of fuzzy control model and queuing model; the dynamic coordination in the auxiliary road: the adaptive system has a high degree of application in this kind of control, which is more used by the ALAIN control system, and the multi system cooperative control aspect. The main application of the control scheme is the joint control of the ramp and the auxiliary road intersection, or through the coordination of the express and entrance control system, from the whole to master the running state of the traffic flow in the express road; the following will be classified and stated.

3.1 Ramp Outlet Direction of Car and Flow Priority

The earliest way to regulate the direction of ramp exit is closing the ramp exit, which is mainly applied to the control of the exit flow in the expressway. However, due to too many disadvantages after closing the ramp, it will cause more traffic problems and has been less applied to the ramp control. At present, the commonly used control means mainly refer to the main road priority [2–4] control idea. The traffic flow in the direction of the ramp exit is regarded as the main road traffic. The idea was first applied to the formulation of the scheduling rules for plane non signal intersection. Zhou [5] established a general model for the inter-passability of two-lane bypass vehicles and pointed out that this model can be applied to the study of highway confluence. Ma [6] used the simulation method to calibrate the main parameters. Considering the large fluctuation of vehicle on the ramp, the intelligent control system is applied in most of the control schemes for real-time control. Among them, the research of fuzzy control system in other fields of traffic has become mature, and has a wide range of application, and it also occupies a certain proportion in the study of expressway exit control. In the design of the control scheme for priority traffic flow in the direction of the expressway exit, Most of the control schemes are based on the queuing model for signal control, and the control scheme is designed based on whether the exit ramp vehicles form a long queue or not. The following two categories of control methods are listed.

3.1.1 Fuzzy Control System Scheme

Fuzzy control is often used to solve complex and uncertain [7] problems because it can simulate the operation mode of human brain. The earliest application of fuzzy control was put forward by L. A. Zadeh in 1965 [8], to solve the uncertain and complex characteristics of the traffic system itself. Fuzzy control has been applied to the traffic system [9–11] by many scholars. In the exportation control of expressway,

Chen and other [12] apply fuzzy algorithm to the design of controller in order to optimize the phase green delay and phase sequence.

The basic expression of fuzzy control is:

$$E = f(S, C, P) \tag{1}$$

The parameters are as follows:

E—system output vector. It can choose the extension of green letter, phase difference or signal cycle;
F—fuzzy system mapping relation;
S—traffic condition parameters. Such as vehicle queue length, total vehicle flow and average speed;
C—other control parameters of the system. Such as interface parameters in multiple controllers;
P—fuzzy system structure parameters, such as fuzzy membership function, fuzzy rules and other.

In order to reduce the error of fuzzy control in large fluctuation of data, the control strategy is to fuzzification some factors affecting the running road vehicles, and then the visual data is changed into different regional range, and the current control scheme is determined according to the region in which the current influence factors are in.

The fuzzy control model has strong robustness and high fault tolerance. It can give good processing results in the face of complex expressway traffic flow condition. However, the process of processing information is simple, which leads to low accuracy and lack of systematicness.

3.1.2 Control Scheme Based on Queuing Model

In order to control the exit ramp priority, some studies take into account the research on the control mode of the exit ramp queuing. By considering whether there is a super long queue at the exit ramp to judge whether the expressway is allowed to take priority.

Among them, Zhu [13] is based on the improved queuing model of the independent entrance of the NEMA. On this basis, the priority control induction algorithm is added to determine whether there is a long queue in the connection road to determine whether the exit ramp can be preferred. Wang and other [14] put forward the integrated control method of the acceptable queuing on the ramp, replacing the detector at the landing point of the ramp by setting the detector at the acceptable line length instead of setting the detector at the landing point of the ramp. When the traffic line at the ramp exit reaches a certain number, the signal state is changed, and the main road passes through the main road.

The queuing model is almost the foundation of the existing research. By judging the queuing conditions of the ramp vehicles for the change of the signal state, how

to judge the location of the exportation ramp detector can make the vehicle release the maximum benefit is the difference in the present research.

3.2 Dynamic Coordination of Auxiliary Roads

The dynamic coordination of auxiliary roads mainly adjusts the traffic in auxiliary roads, thus changing the traffic condition in the expressway exit. By setting up a detection system at the exit of the expressway to sense the queue of vehicles, if it exceeds a certain value, it is necessary to change the signal control at the intersection. Because of the better management effect of adaptive system for the dynamic traffic system, the majority of these categories are in the category, and the MLD hybrid logic dynamic control scheme is more suitable for the traffic flow complex state. The research status of these two control schemes will be summarized below.

3.2.1 Adaptive Control System Scheme

The adaptive algorithm is widely used in the expressway entrance ramp abroad. This kind of algorithm focuses on the flexible adjustment the changing of traffic flow on the ramp, and can adjust the control scheme of the intersection according to the detected data from the detector in real time, so as to achieve the purpose of blocking the traffic in the direction of the exit quickly and reduce the delay of the road network. To alleviate the congestion of the expressway. For example, the adaptive control method of freeway ramp and adjacent intersection proposed by Kwon [15]. But for China's national conditions, the algorithm has also been applied to dredging off ramp traffic conditions.

The ALINEA control model is the most widely used in the adaptive control model. The model is widely applied abroad and the research results are mature. In terms of China's Expressway Control, ALINEA control model has also obtained better theoretical results in Expressway Control [16–18], and designs the adaptive algorithm formula of the auxiliary road intersection as follows:

$$R_{\text{side}}(i+1) = q_{\text{side}}(i) + \varphi(O_s - O_{off}(i)) \tag{2}$$

The parameters are as follows:

$R_{\text{side}}(i+1)$—Vehicle regulation rate at the (i + 1)-th sampling interval; $q_{\text{side}}(i)$—The i-th sampling interval auxiliary channel actually detects the flow through the parking line; φ—Adjustment factor; O_s—Target vehicle occupancy rate for export ramps; $O_{off}(i)$—The time share of the exit vehicle detection at the i-th sampling interval.

In the study of expressway exit control, Qu [19], based on the principle of ALINEA control method [20] adaptive control model algorithm, adopts the signal coordination to the auxiliary road intersection,

Huang [21] analyzed the advantages and disadvantages of the ALINEA model and analyzed the congestion causes of the expressway by using VISSIM software, and then adopted the ALINEA control method to adjust the ramp flow of the exit. The results showed that the model could improve the running efficiency of the vehicle flow.

Besides the ALINEA control system, there is an adaptive algorithm for Wang [22] who use Q-learning method and extension control to control the ramp control. It can adjust the timing scheme of the auxiliary road intersection using Q-learning method, which can automatically adjust the control scheme suitable for the current flow state.

The adaptive control system is widely used in the traffic field because it can dynamically adjust the signal timing scheme to minimize the possibility of vehicle congestion.

3.2.2 CTM Cell Transmission System

The earliest proposed CTM (Cell Transmission model) is Carlos, Daganzo [23]. Because this model can better describe the operational status of traffic flow diversion and confluence. As a result, it has gradually been widely used in transportation network analysis. The initial CTM model has the disadvantage of limiting the length of segments, which leads to the fact that it can not be fitted to the actual traffic state, and its application is limited. Until the CTM model was improved by Munoz and the other [24], then it can be applied to different segment lengths, thus further optimizing the CTM model.

Liu [25] mentioned in his paper that the basic model of CTM is:

$$y_i(k) = \min\{n_{i-1}(k),\, Q_i(k),\, \omega(N_i(k) - n_i(k))/v\} \tag{6}$$

$$n_i(k+1) = n_i(k) + y_i(k) - y_{i+1}(k) \tag{7}$$

The parameters are as follows:

v -Self-owned flow speed; ω—The reverse propagation speed of traffic waves during traffic congestion; $q_{\max}$—Maximum traffic flow; ρ_{i}—Blocking density; $\rho_i(k)$—The traffic density on cell i-th in the k-th period; $n_i(k)$—The number of vehicles on cell i-th in the k-th period; $y_i(k)$—Inflow rate on cells i-th in the k-th period; $Q_i(k)$—The maximum inflow of cells i-th in the k-th period; $N_i(k)$—The maximum capacity of cell i-th in the k-th period.

Yuan and other [26] use the CTM cell transport model to describe the operating state of the main road and the auxiliary road, using the LWR model [27, 28] to simulate the running of the traffic flow, and observe the effect of the auxiliary road signal control on the auxiliary road under different circumstances. Yang and other [29] also applied the CTM to simulate the traffic flow status and referred to the proposed integrated control model for expressway exit ramps and articulated intersections.

When applying the model, it considers whether the road contains bus lanes when connecting ordinary roads, and gives the basic road sections. A reference model was established at the head and tail of the sub-areas and confluences, as well as ordinary roads, and a green light optimization design was performed on the models. Rao [30] applies the improved CTM model to establish the regional control model, and adds the interleaving area to the CTM system to form the CTM weaving area model.

The CTM cell transport model is widely used in traffic control. In the expressway, it is used to simulate the running state of the road traffic, and the effect is good. However, in the simulation process, the model can not fully consider the road shape, for example, the traffic characteristics of the vehicle will change when the road has the turning area.

3.2.3 MLD Hybrid Logic Dynamic Control Scheme

With the continuous improvement of the precision and complexity of the system, when the traditional modeling method has not met the design requirements, Bemporad and Morari proposed a MLD hybrid logic dynamic model, which is used to more accurately solve the control problem of complex systems. As a prediction model, MLD can dynamically adjust the prediction results according to the real-time state changes of reference factors and obtain timely adjustment control scheme.

Wang and other [31, 32] give the MLD model:

$$x(t+1) = Ax(t) + B_1u(t) + B_2\delta(t) + B_3z(t) \tag{3}$$

$$y(t) = Cx(t) + D_1u(t) + D_2\delta(t) + D_3z(t) \tag{4}$$

$$E_2\delta(t) + E_3z(t) \leq E_4x(t) + E_5 \tag{5}$$

The parameters are as follows:

X—System state variables; y—output variable; u—control variable; δ—auxiliary logic variable; z—auxiliary continuous variable.

Wang Li and others applied the MLD model to solve the problem of how to minimize the number of cars queued on the exit ramp. To solve this problem, a MLD hybrid logic dynamic modeling method was established to establish a coordinated calculation model for traffic lights at the exits of the ramp and the adjacent intersections.

The MLD hybrid logic dynamic model has low application in the traffic field, but as a method to predict the state of the hybrid system well, it can provide a better design idea for the expressway control.

The adaptive system is used to collect and simulate the real-time traffic flow data, and the MLD model can predict the running status of the ramp flow, and predict and manage the data through the predicted data. The two systems can regulate and control the ramp of the freeway with complex traffic conditions and fluctuate.

3.3 Multi System Cooperative Control Aspect

In order to ensure the efficiency and safety of the vehicle leaving the expressway, it is necessary to ensure the reasonableness of the management of the ramp and adjacent auxiliary roads, which does not affect the vehicle of the express road as far as possible, but does not have a great influence on the auxiliary road vehicles, and should prevent the collision between the express vehicle and the auxiliary vehicle. Some control schemes select the real-time interaction between the secondary road intersection signal and the expressway ramp control to achieve the purpose of dynamically adjusting the expressway traffic state.

Wang [33] proposed the optimal synergistic model of the ramp control and multi intersection control, which can be used to realize the real-time dynamic coordination of the expressway ramp and the multiple intersection of the pavement, and adjust the vehicle running state flexibly at the ramp of the expressway. Fan and other [34] use the Agent system [35] to consider the entry of Agent to feed the collected data to the regional Agent system, and feed back to the express way by adjusting the traffic flow state of the adjacent intersection, so as to achieve the purpose of dynamically adjusting the running state of the Express vehicle.

The simultaneous adjustment of the ramp and auxiliary road at the express road can effectively avoid the conflict between the expressway vehicles and the auxiliary road vehicles, and the benefits of the auxiliary road vehicles can not be greatly damaged.

4 Personal Review

At present, the control of expressway exit in China mostly solves the problem of congestion at the exit of Expressway by adjusting the traffic flow comes from the auxiliary road or auxiliary road nearest to the intersection. The coordinated control of the auxiliary road is the high application degree control method of the express road exit control. In our country, some express road exports have obtained better control results. Because they do not need complex algorithm design and higher equipment cost. So the application degree is high.

In the control system which is applied to the auxiliary road and adjacent intersection, the two control schemes, the fuzzy control system and the adaptive system, have already had many achievements in other fields of traffic, and the theoretical research is more mature. Therefore, we can have a comprehensive grasp of their advantages and disadvantages in the control of expressway exit, as well as the applicable conditions. Especially for the complex traffic condition of expressway exit, a good plan can be chosen.

However, it is necessary to pay attention to the adjustment for the traffic flow of the auxiliary road and adjacent intersection, so that the interests of other roads will be sacrificed. With the continuous increase of road vehicles, the control ability of

this kind of control is limited and will not guarantee the traffic efficiency of vehicles. When the number of vehicles exceeds a certain limit, it will exceed the ability of the intersection to regulate and control, and this will affect vehicle traffic.

Therefore, it is possible to expand the scope of the ramp control step by step, not only to control the adjacent auxiliary road intersection of the ramp exit, but to extend the control range from two to three adjacent intersections, so as to avoid the possible restrictions on the control of a single intersection.

With the continuous development of intelligent traffic, the feasibility of the synergetic control of the exit and entrance is continuously improved. The stability and order of the expressway operation can be maintained effectively from the overall regulation of the vehicle flow. Although a large number of detection devices and complex intelligent control systems are still needed. The control effect will be better than the "decompression" of the expressway through the auxiliary road coordination.

5 Prospects

With the deepening of the research on the control of expressway exit, the traffic factors that need to be considered are increasing. This paper will give some thinking and Prospect on the development of the study on ramp control in the future.

(1) The control of expressway exit is mostly focused on the microscopic study of vehicle running state, such as the CTM cell transmission model. The two models should be organically combined to establish a model system that can connect the whole with the microcosmic, which can effectively improve the efficiency of Expressway Ramp Traffic.
(2) With the development of intelligent traffic, cooperative control is the most ideal control state of the future Expressway Control. Then, it is necessary to consider how to ensure the accuracy of the data after the system processing, the accuracy of the system control scheme, and the flow standard when the control is needed. All of these need to find reasonable control standards.
(3) The control of the auxiliary road intersection is the main solution for the current ramp congestion problem, but this method will also bring some traffic pressure to the auxiliary road and the surrounding road. How to ensure the smooth passage of ramp exit on the basis of other roads does not cause obstacles, is the need for further study. Moreover, with the increasing number of road traffic and the increase of auxiliary road traffic, it is necessary to consider whether the auxiliary road will be controlled.
(4) At present, many researches on the exit of expressway exits refer to the way of entrance control. Because the flow status of the exit and entrance is different in some cases, the control algorithm which is more suitable for the flow characteristics on the ramp exit should be considered.

6 Conclusion

In this paper, the existing control methods of expressway exit are divided into three types: export ramp flow priority, exit ramp flow priority, auxiliary road dynamic regulation and multi system cooperative control, and explain the individual viewpoint and prospect for future development. Although some theoretical achievements have been achieved, the application effect still needs further observation. Considering the scale of urban expressway development in the future, effective congestion control should be adopted as soon as possible to prevent congestion.

References

1. Li S (2007) Research on the optimization method of control parameters for the expressway exit on the urban expressway. Jilin University
2. Yu W (2004) Simulation study on the traffic capacity of urban main roads at priority control intersections. Wuhan Municipal People's government and urban transport planning academic committee of China Urban Planning Society. Urban transportation development model transformation and innovation: China urban transport planning 2011 annual conference and the 25th academic symposium proceedings. Wuhan Municipal People's Government, Urban planning council of China Urban Planning Society, 9
3. Ma D, Wang D, Yang X, Chen S (2010) Critical flow based on signal setting at a limited priority intersection. J SE Univ (Nat Sci Ed) 40(04):860–865
4. Guo X, Yuan Z (2006) Analysis of traffic capacity of main road priority control intersection. J Shandong Univ Technol (Nat Sci) 06:89–92
5. Zhou W (1990) Discussion on the ability to intervene vehicles at the intersections of non-signalized intersections. J Chang'an Univ (Nat Sci Ed) 03:21–27
6. Ma X, Wang W (1997) Application of simulation method in the ability of main road priority intersection. J SE Univ 5:73–77
7. Wu Z, Yin L (2008) Analysis of principle and application of fuzzy control system. J Gansu Lianhe Univ (Nat Sci Ed) 01:91–94 + 98
8. Zadeh LA (1965) Fuzzy sets*. Inf Control 8:338–353
9. Zhang J, Zhang D, Chen C (2005) Comprehensive fuzzy control of urban intersections. J Hefei Univ Technol 28(3):229–233
10. Zhou Z, Wang Y, Wang X (2009) A fuzzy control method for multi-phase intersections. Comput Eng 35(1):168–172
11. Zheng J, Chen H, Dong D (2006) Design and simulation of urban freeway on-ramp fuzzy controller. J Highw Transp Res Dev 12:133–136
12. Chen X, Wang N (2013) Fuzzy control of timing matching at an intersection of expressway. J Kunming Univ Sci Technol 38(01):45–49
13. Zhu H-g (2010) Study on the coordination control problem of expressway exits and articulated roads intersections. Traffic Stand 14:119–122
14. Wang H (2015) Research on the integrated control of urban expressway exit ramps and articulation intersections. J Shandong Jiaotong Univ 23(04):17–22
15. Kwon E, Ambadipudi RP, Bieniek J (2003) Adaptive coordination of ramp meter and intersection signal for optimal management of freeway corridor
16. Liu X (2013) Intelligent learning control method for expressway traffic entrance ramp. Qingdao University of Science and Technology
17. Zheng F, Du Y, Sun L (2009) On-ramp single-point dynamic control based on ALINEA algorithm. J Tongji Univ (Nat Sci) 37(06):766–771

18. Chi R, Hou Z, Yu S (2008) Nonparametric adaptive iterative learning control for freeway ramps. Control Theor Appl 25(06):1011–1015
19. Qu Z, Wang L, Jin S (2009) Simulation research on link control of expressway exit ramp. Traffic Inf Secur 27(02):43–47
20. Papageorgiou M, Hadj-Salemand H, Blosseville JM (1991) ALINEA: a local feedback control law for on-ramp metering. Transp Res Rec 1320:58–64
21. Huang A-q (2011) Research on traffic flow analysis and control method of urban expressway main and auxiliary roads. Beijing Jiaotong University
22. Wang X (2011) Research on expressway exit control method. North China University of Technology
23. Daganzo CF (1994) The cell transmission model. Part I: a simple dynamic representation of highway traffic. Transp Res Part B 28(4):269–287
24. Munoz L, Sun X, Horowitz R (2008) Traffic density estimation with the cell transmission model
25. Liu Y (2011) Research on traffic control and optimization of intersections based on cellular transmission model. Beijing Jiaotong University
26. Yuan C, Li H (2009) Research on the signal control strategy of auxiliary road under the condition of different expressway exit design capacity. J Highw Transp Res Dev 26(01):139–143
27. Lighthill MJ, Whitham GB (1955) On kinematic waves i: flood movement in long rivers, II: a theory of traffic flow on long crowded roads. Proc R Soc A229:281–345
28. Richards PL (1956) Shock waves on the highway. Oper Res 4:42–51
29. Yang X, Han Y, Fu Q (2009) Integrated control model of ramps and articulated intersections on urban expressway. J Transp Eng 9(02):110–115
30. Rao J (2014) Research on modeling and control of urban expressway exit ramp region. Zhejiang University
31. Wang L, Li Z, Xiu W, Liu X (2014) Coordinated control of expressway exit and auxiliary road based on MLD-MPC. Control Eng 21(04):487–490 + 495
32. Xiu W (2013) Research on controlling technology of urban expressway exit signal based on MLD. North China University of Technology
33. Wang S (2014) Research on key technologies of collaborative control of traffic signals in expressway ramp connection. South China University of Technology
34. Fan L, Chen Y, Li Z (2009) Research on coordinated control of urban expressway based on multi-agent. Traffic Inf Secur 27(04):30–34
35. Yang H, Yan Y, Liu D (1999) Agent: characteristics and classification. Comput Sci 09:30–34

Study on Community Detection of Shipping Network Based on Modularity

Xuejun Feng, He Jiang and Liu-peng Jiang

Abstract Shipping network is a kind of typical complex network, the network structure is one of its important features. This paper takes the research of community structure of shipping network as the object, constructs the Newman fast algorithm based on modularity, and choose "The twenty-first Century Maritime Silk Road" shipping network as the case, which is unweighted and undirected shipping network, and composed of 453 ports and 3444 edges. From the perspective of shipping network connectivity, the Newman fast algorithm is used to calculate "The twenty-first Century Maritime Silk Road" shipping network. The structural properties of this shipping network can be obtained. There is only one core community in this shipping network, which is leader community, and consists of 173 ports. Their degree follows the power-law distribution. Others are non-core communities. It shows that the "The twenty-first Century Maritime Silk Road" container shipping network owns huge community structure with core nodes. The conclusion of the research is a reference to the relationship between "The twenty-first Century Maritime Silk Road" shipping network and the ports along its line.

Keywords "The twenty-first century Maritime Silk Road" shipping network · Complex network · Community detection · Network structure

1 Introduction

With the rapid progress of global economic integration, shipping network plays a bigger and bigger role. The shipping network, which is a typical complex network [1], is mainly composed of the ports (vertices) and the routes (edges). Due to the large number of elements and scale in the shipping network, many scholars divided it into port subsystem and route subsystem to carry out separate analysis [2–5], but lack of integration. This paper analyzes the flow of containers on the shipping network from

X. Feng · H. Jiang · L. Jiang (✉)
College of Harbor Coastal and Offshore Engineering,
Hohai University, Nanjing 210098, China
e-mail: jsjlp@hhu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_28

complex network perspective, and explore the property and structure of shipping network.

As the frontier of the research of complexity theory, complex network theory is produced by the fusion of many disciplines such as natural science and social science. As a new research tool, scholars pay more and more attention to it. After Watts and Strogtz put forward the WS small-world network model [6, 7], and Barabasi and Albert proposed the BA scale-free network model [8], the research results of complex network show explosive growth, which makes the research on complex networks is no longer confined to classical physics and statistical physics. Understanding the properties of complex networks is becoming a research hotspot from the perspective of the whole system. The application of complex network theory to the port shipping system is relatively late, and the related results are less, which mainly focus on the following aspects. ① The research on the complexity properties of shipping network: By means of statistical analysis, Wu et al. [9] proved that the container shipping network has the properties of scale-free and small-world. Mu et al. [10] carried out computer simulation of the realistic route data, and analyzed its topological properties by using the complex network theory, which verified that shipping network has the property of rich club. Chen et al. [11] verified that the maritime silk road shipping network in Southeast Asia conforms to the properties of complex network based on the complex theory. Zhang et al. [12], Ye et al. [13], Wei et al. [14] studied the optimization and scheduling of routes. ② The research on the structural features of shipping network: Wilmsmeier et al. [2] analyzed the spatial evolution of the ports in Latin America and the Caribbean and obtained the spatial layout and its properties of ports in the container shipping network. Feng et al. [3, 4], Jian et al. [5] obtained the spatial structure and evolution trend of port by analyzing the developing rules of port system from the perspective of port group or route. ③ Analysis of shipping network structure from the perspective of network: Ducruet [15] conducted an in-depth research on the periodicity and dynamics of global shipping network. Based on the global maritime flow in 1977–2008, the method of complex network was adopted to obtain the clustering properties of network growth and the distribution properties of traffic volume [16]. Ducruet et al. [17] analyzed the Chinese global shipping connectivity from the perspective of shipping dynamics between 1890 and 2016, and emphasized the influence of technology, economy and politics on the distribution of internal and external connectivity.

In the beginning of this century, Girvan and Newman discovered and studied a topological structural property of complex networks—community structure [18]: the connections among communities in complex network are relatively sparse and the connections within the communities are relatively dense, which makes community structure becomes one of the important properties of complex networks. After that, Bichot et al. [19], Traff [20], Zamprogno and Amaral [21], Benlic and Hao [22] studied the partitioning method of complex network. Guo et al. [23], Liu et al. [24] has solved the specific problems in the actual network based on existing network partitioning methods. These researches and applications make community structure get more and more attention as an important symbol of complex network properties.

Overall, the existing research results mainly focus on the properties of shipping network. They demonstrate the small-world property and scale-free property of shipping network, and calculate the general index of shipping network such as degree, degree distribution, centrality and betweenness. The research on the community structure of shipping network is less fruitful, some unique properties of shipping network can not be reflected in the network model such as the concepts of hub port and port group. Aiming at the deficiency of existing research, this paper studies the structure of shipping network based on the method of community detection from the perspective of complex network.

2 Elements of Shipping Network and Data Acquisition

2.1 General Complexity Properties of Complex Network

The topological structure of the real complex network is not completely regular, nor randomly generates, but a complex network structure which is formed by local partial subgraphs with partial random connections. With the breakthrough of complex network theory, the method system of network analysis is continuously enriched and perfected. At present, the mainstream properties of complex network topology and its characteristic condition are shown in Table 1. The main properties of the network can be determined by analyzing the main characteristics of the network, such as degree distribution, average path length and clustering coefficient.

In the characteristic condition corresponding to the topological property, the degree of node is the number of edges that directly connected to the node, which reflects the level of connectivity of node in complex network. By calculating the degree of node, the degree distribution of all nodes can be understood in the network. Average path length is the average value of the minimum number of edges that are needed to connect any two nodes in the network. The average path length of network can be obtained by averaging the distance among all node pairs. In general, average path length is relatively small in complex network. The clustering coefficient

Table 1 The common topological properties of network and its main characteristic condition

Serial number	The topological property of complex network	Characteristic condition
1	Small-world	The smaller average path length, the higher clustering coefficient
2	Scale-free	The degree distribution conforms to the power-law distribution
3	Stochastic	The degree distribution conforms to Poisson distribution

represents the probability that the nodes connected to the same node are interconnected in complex network. The clustering coefficient of a node is equal to the ratio of the number of edges connected to all other nodes to the maximum number of edges that may exist.

2.2 Analysis and Acquisition of the Elements

The shipping network is mainly composed of ports and routes. Port are the node in the network, and route connecting port and port is edge. Shipping networks can be defined as $G(V, E)$. $V = \{v_i | i = 1, 2, \ldots, n\}$ is a collection of nodes. n is the number of ports. $E = \left\{e_{ij} | i, j = 1, 2, \ldots, m, i \neq j\right\}$ are the edges of shipping network. m is the number of routes. This network is unweighted and undirected network.

The data in this paper is from the container liner schedule of China International Shipping Network (www.chinashipping.com). The ports within the scope of the research are selected. The range of the edges is all the liner routes within the scope of the research. The crawler software based on JAVA is used to convert the data to the degree matrix of network.

2.3 Community Structure of Shipping Network

Community is a collection of the units with certain common characteristics. Node set in complex network can be classified into different subsets according to their properties. The primary target of community detection is to discover node groups with similar properties by topological properties of network. Analyzing the community structure in complex network helps to understand its internal structure and the function of each community for the entire network. As a kind of complex network, shipping network also has community structure. In addition to the topological properties of common complex network, the shipping network has its peculiar properties, which are influenced by the spatial distance, politics, culture, trade and other factors. Dividing of the community structure shipping network and analyzing its structural characteristics provide a new perspective for the analysis of trade among the ports, and it is convenient to formulate policies, adjust trade strategies and allocate resources rationally.

3 Method of Community Detection Based on Modularity

3.1 Measurement of Community Structure—Modularity

Currently, the general accepted measurement of community structure is modularity [25] defined by Newman and Girvan. Modularity is the expectation of the difference between the ratio of the edges connecting the internal nodes of community and the ratio of the edges of any node connected to the network. Function Q can be used to quantitatively describe the modular level of community detection.

Suppose that network has been divided into communities, the number of communities is k. The k*k-dimensional symmetric matrix is defined as $e = (e_{ij})$. The matrix element e_{ij} represents the ratio of the number of edges connected between community i and community j to the total number of edges in the network. Suppose that $Tre = \sum_i e_{ii}$ is the sum of the elements in the diagonal of the matrix, which represents the proportion of the number of edges connected among the internal nodes of community in the total number of edges in the network. Suppose that a_i is the sum of the elements in each row or each column, which represents the proportion of the edges connected to the nodes in community i in all the edges. Therefore, function Q can be expressed as.

$$Q = \sum_i (e_{ii} - a_i^2) = Tre - \|e^2\| \tag{1}$$

$\|e^2\|$ represents the sum of all the elements in matrix e^2. If the expected value of the ratio of the inner edges of community is not greater than the expected value of the ratio of the edges of any connected nodes, $Q \leq 0$. The upper limit of Q is $Q = 1$. Q is closer to 1, community structure is more obvious, and better divided. In the actual network, the value of Q is usually between 0.3 and 0.7. In the process of dividing the network, the corresponding modularity (the value of Q) of each kind of division is calculated, and one or several kind of divisions of the maximum value of Q will be found, which is the optimal division of community structure.

3.2 Newman Fast Algorithm Based on Modularity

The classical GN algorithm [18] was proposed by two scholars, Girvan and Newman. GN algorithm is a kind of algorithm for splitting community. It is about community detection based on the optimization of modularity. The edges are removed from the network according to descending order of betweenness, until the whole network is divided into communities. On account of the high time complexity in calculating betweenness, and betweenness needs to be calculated before each partition in network, which greatly increases the complexity of algorithm. In order to improve the time complexity of GN algorithm, Newman proposed the Newman fast algorithm

[26] based on GN algorithm. Newman fast algorithm is a kind of greedy algorithm to find the maximum value of modularity, and also belongs to the aggregation algorithm for adding edges to the network. The core idea of the algorithm is a bottom-up hierarchical clustering process. This method can be used to detect community structure by maximizing modularity, which greatly reduces the time complexity and can be used to analyze the complex network with millions of nodes. The specific algorithm process is as follows:

(1) First of all, the complex network with n nodes is regarded as n communities. The initial case is as follows:

$$e_{ii} = \begin{cases} 1/2m & \text{There is connection between node } i \text{ and node } j \\ 0 & \text{Otherwise} \end{cases} \tag{2}$$

$$a_i = k_i/2m \tag{3}$$

where in: k_i is the degree of node i. m is the total number of edges contained in complex network.

(2) According to the principle of greedy algorithm, two connected communities are merged in turn along the direction that makes Q increase or decrease fastest. Then calculate the increment of Q:

$$\Delta Q = e_{ij} + e_{ji} - 2a_i a_j = 2(e_{ij} - a_i a_j) \tag{4}$$

After the merging, update the corresponding elements and add the rows and columns of community i and community j.

(3) Repeat the previous step until each node merges into a community.

4 Case Analysis

The paper selects "The twenty-first Century Maritime Silk Road" shipping network as the research object. The time range is from May 2017 to September 2017. And it is composed of 453 ports and 3444 edges. The geographical distribution of ports is shown in Fig. 1.

The complex network consisting of 453 nodes and interconnected edges is divided into 225 communities by using the community structure division method in Sect. 2.2. These communities are divided into three categories according to the number of nodes and their topological properties. The results of community division are shown in Table 2.

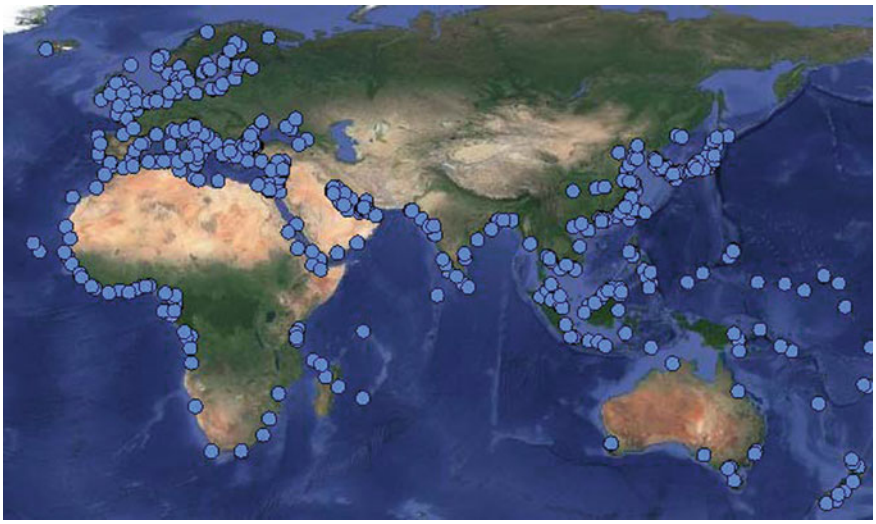


Fig. 1 Geographical distribution of ports

Table 2 The results for community division

Category number	The number of communities in this category	The number of ports in this category	The number (number range) of nodes within each community	The range of degree
1	1	173	173	8–207
2	25	81	2–7	1–17
3	199	199	1	1–27

As shown in Table 2, "The twenty-first Century Maritime Silk Road" shipping network has giant community structure, and there is only one giant community (The first category of community) in it. The first category of community consists of 173 ports. The ports with the top 20% of ranking of degrees in "The twenty-first Century Maritime Silk Road" shipping network are located in this community. The frequency distribution of degrees in the first category of community is shown in Fig. 2.

From Fig. 3, the degree of the first category of community are close to the power-law distribution, which meets the leader community. There are several core ports in the first category of community. The second category of community are generally small with lower degrees and obvious geographical distribution, which is mainly distributed in Europe. It can be seen that European countries have close trade links with each other under the influence of the European Union, but there is still small communities independent of the core community.

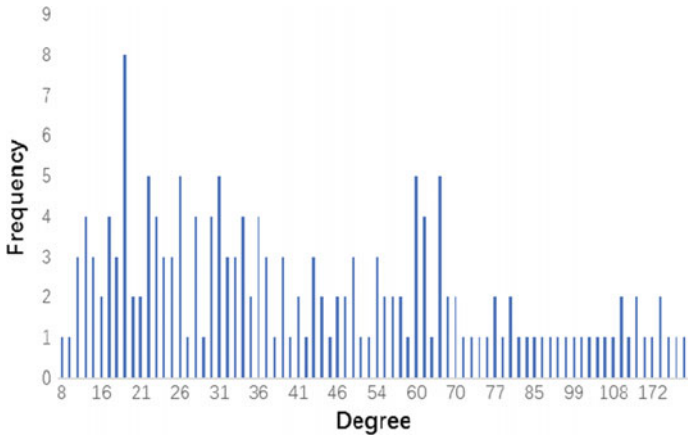


Fig. 2 Distribution of degree for the first community

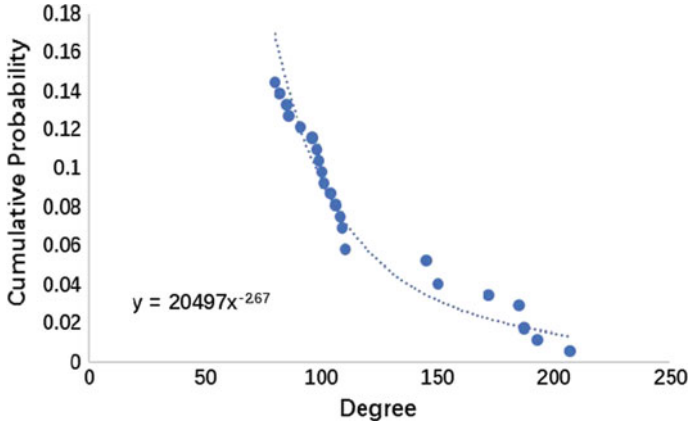


Fig. 3 Proportion diagram of cumulative probability of nodes in the top 22 for the first community

5 Conclusion

This paper takes the research of community structure of shipping network as the object, constructs the Newman fast algorithm based on modularity, choose "The twenty-first Century Maritime Silk Road" shipping network as the case, and divide the shipping network which consists of 453 ports and routes. The results show that "The twenty-first Century Maritime Silk Road" container shipping network has a giant community, which belongs to the leader community. It shows that "The twenty-first Century Maritime Silk Road" container shipping network is a huge community structure with core nodes. This conclusion can provide the technical support for the development of "The twenty-first Century Maritime Silk Road" container shipping

network. The research of the interaction mechanism between the core nodes and the whole shipping network will be the next research direction of this paper.

References

1. Wang N, Wu N, Dong L-l, Yan H-k, Di W (2016) A study of the temporal robustness of the growing global container shipping network. Sci Rep 6:1–10
2. Wilmsmeier G, Monios J, Pérez-Salas G (2014) Port system evolution—the case of Latin America and the Caribbean. J Transp Geogr 39(39):208–221
3. Feng X, Wang W, Jiang L (2008) Optimization model and algorithm of port cluster system. J Traffic Transp Eng 8(3):77–81
4. Jiang L, Feng X, Wang W (2011) Port—industry—city composite system coordination model. J Econ Water Resour 29(1):11–14
5. Jian L, Li D, Liu L (2012) Research on the evolution law of China container port system. Econ Geogr 32(12):91–96
6. Wang QS, Xiao F (2011) Efficient routing on lager road networks using hierarchical communities. IEEE Trans Intell Transp Syst 12(3):132–141
7. Watts DJ, Strogatz SH (1998) Collective dynamic of 'small-world' networks. Nature 393(6684):440–443
8. Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–512
9. Wu P, Deng S, Tian W (2008) Research on topology character of container shipping network. J Wuhan Univ Technol (Transp Sci Eng) 04:665–668
10. Mu , Chen Y, Yang M, Li T (2009) Topological features of liner shipping network. J Dalian Marit Univ 35(02):34–37
11. Chen F, Hu Z (2016) Analysis of shipping network of Southeast-Asian route along maritime Silk Road based on complex networks. J Shanghai Univ (Nat Sci) 22(06):804–812
12. Zhang Y, Yang H, Ji M et al (2016) Liner service network design with speed optimization. J Transp Syst Eng Inf Technol 16(5):219–226
13. Ye D, Huang Y, Hu J (2015) Study on fleet deployment question for liners with low carbon pattern. J Central China Normal Univ (Nat Sci) 49(2):322–326
14. Wei Z, Xie X, Wei M (2016) Container ship speed optimization under influence of wind and wave. J Transp Syst Eng Inf Technol 16(3):154–160
15. Ducruet C (2015) Maritime networks. Spatial structures and time dynamics
16. Ducruet C (2017) Multilayer dynamics of complex spatial networks: the case of global maritime flows (1977–2008). J Transp Geogr 60:47–58
17. China's global shipping connectivity: internal and external dynamics in the contemporary era (1890–2016). Chin Geogr Sci 2:202–216 (2018)
18. Girvan M, Newman ME (2001) Community structure in social and biological networks. Proc Natl Acad Sci USA 99(12):7821
19. Bichot CE (2007) Application of fusion-fission to the multi-way graph partitioning problem. Lect Notes Comput Sci 4967:698–707
20. Traff JL (2006) Direct graph k-partitioning with a Kernighan-Lin like heuristic. Oper Res Lett 34(6):621–629
21. Zamprogno R, Amaral ARS (2007) An efficient approach for large scale graph partitioning. J Comb Optim 13(4):289–320
22. Benlic U, Hao JK (2011) An effective multilevel tabu search approach for balanced graph partitioning. Comput Oper Res 38(7):1066–1075
23. Guo L, Zhu Y (2005) Application of social network analysis on structure and interpersonal character of sports team. China Sport Sci Technol 41(5):10–13

24. Liu Z, Li S, Lin D et al (2009) Blog community discovery based on PCM clustering algorithm. J Xiamen Univ (Nat Sci) 48(4):508–513
25. Newman MEJ (2004) Detecting community structure in networks. Eur Phys J B 38(2):321–330
26. Newman MEJ (2004) Fast algorithm for detecting community structure in networks. Phys Rev E Stat Nonlinear Soft Matter Phys 69(6 Pt 2):066133

Research on Factors Affecting the Effect of Chinese Port Transformation and Upgrading

Xuejun Feng, Jiaojiao Wang and Liupeng Jiang

Abstract Chinese ports are in a critical period of transforming development methods. It is necessary to scientifically realize the transformation and upgrading of ports. However, the factors affecting the port transformation and upgrading are complex, multi-layered and difficult to quantify. All kinds of influencing factors and their relationships are considered comprehensively and a structural equation model for the influencing factors of port transformation and upgrading is established in this article, based on the structural equation model (SEM). The effect of each influencing factor on port transformation and upgrading is quantitatively analyzed, and then the key influencing factors are identified. This was accomplished by conducting the data sampling and the questionnaire survey. This article provides decision-making basis for port transformation and upgrading, which is of great significance.

Keywords Chinese ports · Transformation and upgrading · Factors · SEM

1 Introduction

As a national important infrastructure and an important hub of integrated transportation system, port is not only the junction point of domestic market and international market, but also the intersection of domestic and international economy. It plays an important role in the national economy and social development. With entering a 'new normal' economy, coastal ports are facing increasingly fierce competition of homogenization along with slower growth in throughput. It is an urgent requirement for the sustainable development of Chinese ports to promote the transformation and upgrading. However, the port transformation and upgrading process is influenced and integrated by many factors, the factors are difficult to grasp, which leads to the lack of quantitative analysis of the research on the transformation and upgrading of ports. Therefore, it is the foundation and difficulty to study the impact of different factors and determine the key influencing factors of port transformation and upgrading.

X. Feng · J. Wang · L. Jiang (✉)
College of Harbour, Coastal and Offshore Engineering, Hohai University, Nanjing 210098, China
e-mail: jsjlp@hhu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_29

The concept of port transformation and upgrading has typical Chinese characteristics. It is a scientific judgment made by Chinese government according to the requirements of the international and domestic economic development to the port industry, which has profound connotation and extensive extension. Because of its distinct characteristics of the times, the relevant research results are relatively scarce. Foreign scholars such as Langen [1], comparing the competitiveness of the lower Mississippi River port group and the Rotterdam port group, pointed out clearly the importance of establishing the port group cooperation mechanism to the competitiveness, and improving the competitiveness of the port group is the key path for the regional port transformation and upgrading; Mccarthy [2] described the construction and development process of the port of Rotterdam since the second world war, and summarized the significance of the transformation and upgrading in the development process of European ports. Kim [3] evaluated the success of port transformation and upgrading by evaluating port performance. Domestic scholars such as Wang [4], Li [5] and Zhang [6] analyzed the situation and existing problems of Chinese port transformation and upgrading in combination with the background of 'One Belt and One Road' strategy, and emphasized that the government should play a role in macro and micro level and put forward relevant suggestions to scientifically realize the transformation and upgrading of ports. There are many scholars had put forward the development strategy of port transformation and upgrading according to local conditions based on the current situation analysis of ports such as Dalian Port [7], Qingdao Port [8] and Hong Kong Port [9]. Overall, the research on port transformation and upgrading at home and abroad mainly focuses on the connotation, model and strategy, and the research on the influencing factors of port transformation and upgrading is insufficient. The reason may be that: on the one hand, the extension of the concept of port transformation and upgrading involves the port infrastructure, the management environment, the port management, the government policy and financial support, etc. Their relationships are complex, and it is difficult to reflect the individual indicators; on the other hand, most of the above indicators cannot be measured accurately and directly, which means that it is difficult to establish a quantifiable evaluation standard.

Structural Equation Model (SEM) is a social scientific statistical method based on the covariance matrix of variables to analyze the relationship between variables, it is also called covariance structure analysis [10]. It can solve the problem of variables that cannot be observed directly, at the same time, it can deal with multiple dependent variables, and allow the independent variables and the dependent variables to contain measurement errors, which make up for the shortcomings of the traditional statistical methods [11]. SEM has been favored by scholars all over the world and widely used in correlation analysis, evaluation index system construction and evaluation of satisfaction. Krishnakumar [12] explored the interaction between children's various abilities by using the structural equation model; Noyan [13] studied the structural determinants of customer satisfaction by building a model of loyalty structural equation. Qi [14] taking transportation industry as an example, defined eight social expectation themes of performance appraisal including the responsibility governance, the economy development and others.

To sum up, most of the current research achievements of port transformation and upgrading are at the macro level, and it is short of the mechanism analysis model and related analysis tools of port transformation and upgrading. Therefore, the key factors of the port transformation and upgrading are difficult to identify, and the effect of each factor is difficult to quantify. The structural equation model has advantages in dealing with multiple variables and revealing the relationship between variables and so on. Hence, it is appropriate to use SEM to identify the key factors and verify the influence degree of each factor of the port transformation and upgrading.

The structural equation model of the influencing factors of the port transformation and upgrading is established in this article, and the key influencing factors are identified through data statistics and analysis, which is of theoretical and practical significance.

2 Structural Equation Model

There are two basic models in the structural equation model: Measurement Model and Structural Model.

The expression of Measurement Model is as follows:

$$\begin{cases} X = \Lambda_X \xi + \delta \\ Y = \Lambda_Y \eta + \varepsilon \end{cases} \tag{1}$$

In the equation,

X vector $(q \times 1)$, which is the observed variable of the latent variable ξ;
Y vector $(p \times 1)$, which is the observed variable of the latent variable η;
Λ_X matrix $(p \times m)$, which is the factor loading of the observed variable X;
Λ_Y matrix $(q \times n)$, which is the factor loading of the observed variable Y;
ξ vector $(n \times 1)$, the independent variable;
η vector $(m \times 1)$, the dependent variable;
δ vector $(q \times 1)$, the measuring error of X;
ε vector $(p \times 1)$, the measuring error of Y.

The expression of the structural model is as follows:

$$\eta = B\eta + \Gamma\xi + \zeta \tag{2}$$

In the equation,

B matrix $(m \times m)$, the coefficient of η in the structure equation;
Γ matrix $(m \times n)$, the coefficient of ξ in the structure equation;
ζ vector $(m \times 1)$, the random error of the structural equation.

The relationship between the potential variables and the observed variables is represented in the measurement model. The structural model is a description of the causal relationship model between potential variables. In general, a structural equation model consists of several measurement models and a structural model.

3 Influencing Factors Analysis

3.1 Influencing Factors of Transformation and Upgrading of Ports

In order to take a more comprehensive consideration of the factors affecting the transformation and upgrading of ports, the influencing factors are divided into two dimensions in this paper, the 'internal factors' and the 'external factors', which is on the basis of the analysis of the existing research literature [15–22], and the related experts, for the more comprehensive consideration of the port transformation and upgrading. The 'internal factors' mainly refer to the existing development level of ports and the management of ports, while the 'external factors' mainly refer to the management environment of ports, the support of government policies and the support of the financial organization.

3.1.1 Internal Factors

The existing development level of port is considered as one of the internal factors which affect the transformation and upgrading of ports. It includes the port infrastructures, the logistics service capacity, the port collecting and distributing transport condition, the port service level, the port technical development level and the port capital strength. First of all, the port infrastructure mainly refers to the terminal berths, operating machineries, yards, warehouses, channels and anchorages, which directly affect the overall operation of the port. Secondly, the port logistics service capacity is mainly reflected in port logistics scale and port logistics operation capacity. Thirdly, the port collecting and distributing transport condition mainly refers to the development level of the transportation system of railway, road and waterway and the intensity of route. Next, the ship time in port, the efficiency of pilotage service, the timely arrival rate of the goods and the safety and satisfaction of the port are considerations of the port service level. Moreover, the port technical development level mainly refers to the existing informatization level, intelligent level and technical research ability. And the last one, the port capital strength is considered from the aspects of capital strength, fixed assets investment and number of employees.

The other internal factor is the port operation and management, which mainly includes the port management level, the diversification of port operation and the extension of the port industry chain. The port management level is considered from

the aspects of port management team ability, strategic vision, talent reserve and quality. The diversification of port operation refers to the diversified development related to the main business, such as expanding shipping finance, insurance, brokerage, consulting and other businesses, in addition to the traditional role of port operation. The extension of the port industry chain refers to the extension of the port industry chain and the additional capacity of production value.

3.1.2 External Factors

The port operating environment is firstly considered as the external factor of port transformation and upgrading, which mainly includes the market environment of port and shipping industry, and the economic condition of the port hinterland. The market environment of port and shipping industry is an important factor that affects the development of the entire port industry, refers to the authoritative index like the Baltic Dry Index (BDI), the Belt and Road shipping index and Chinese port comprehensive competitiveness index, etc. The economic conditions of the port hinterland can be reflected from the gross national product (GDP) of the hinterland of the port. In addition, the foreign trade of the port hinterland also be considered.

Secondly, the support of government policies is also considered as the external factor, which mainly are the government's fiscal and tax preferential policy for ports, the related policies of transformation and upgrading of support and the government's efforts to introduce the professional talents.

Financial institution support is also an indispensable external factor. Financial organization support mainly considers the convenience of ports to create funds and the support of banks for port loans.

3.2 Evaluation Factors of the Effect of Port Transformation and Upgrading

The port enterprises are the main body of port transformation and upgrading. There are two characteristics of the port enterprises in the social economy: one is the enterprise natures and the other is the social public welfare [21].

Considering the characteristics of port enterprises, 'economy' and 'profitability' are the two characteristics of the port. Therefore, the economic benefit is one of the factors to evaluate the effect of port transformation and upgrading. It can be inspected from two aspects of port output value and port throughput.

Considering the social public welfare of the port, the resources, such as funds and talents, can be value-added and reasonably configured through the exchange window of the port. That is, the port enterprises should make a contribution to the development of the economic hinterland at the same time of themselves are profitable, what's more, provide relevant employment opportunities for the society, resulting in

a huge social benefit. Therefore, social benefit is one of the evaluation factors of port transformation and upgrading.

'Green port' is also one of the objectives of port transformation and upgrading. While evaluating the economic and social benefits of the port transformation and upgrading, we should also consider the sustainable development of the environment, that is, environmental benefit. It refers to whether the port is concerned with the current development and seek benefits, regardless of future generations. It can be evaluated in terms of port resource utilization as well as the production and discharge of pollutants.

In conclusion, the influencing factors and evaluation factors of the transformation and upgrading of ports are listed as shown in Table 1.

Table 1 The influencing factors and evaluation factors of the transformation and upgrading of ports

Dimension	Factor	Observed indicator
A1 Internal factors	B1 Existing development level of ports	C1 Port infrastructure
		C2 Logistics service capacity
		C3 Port collecting and distributing transport condition
		C4 Port service level
		C5 Port technology development level
		C6 Port capital strength
	B2 Management of ports	C7 Port management level
		C8 Diversification of port operation
		C9 Extension of the port industry chain
A2 External factors	B3 Management environment of ports	C10 Market environment of port and shipping industry
		C11 Economic situation in port hinterland
	B4 Support of government policies	C12 Fiscal and tax incentives
		C13 Special support policy
		C14 Port industry talent introduction policy
	B5 Support of the financial organization	C15 Financial convenience
		C16 Strength of bank loans

(continued)

Table 1 (continued)

Dimension	Factor	Observed indicator
A3 Evaluation factors	B6 Economic benefit	C17 Port output
		C18 Port throughput
	B7 Social benefit	C19 Solve the employment
		C20 Contribution to hinterland economy
	B8 Environmental benefit	C21 Resource utilization
		C22 Pollutant generation and discharge

4 Model Construction and Questionnaire Analysis

4.1 Model Hypothesis

Based on the analysis of the influence factors of port transformation and upgrading and its effect evaluation factors, the following hypothesis are proposed:

H1: the existing development level of ports has significant positive influence on the port transformation and upgrading. That is, the higher the existing base level of the port, the better economic, social and environmental benefits brought by the transformation and upgrading.
H2: the port operation and management have significant positive influence on port transformation and upgrading. That is, the higher the port operation and management level, the economic benefits, social benefits and environmental benefits brought by the transformation and upgrading are better.
H3: the port economic operating environment has significant positive influence on port transformation and upgrading. That is, the better the port economic operation environment, the economic benefits, social benefits and environmental benefits brought by the transformation and upgrading are better.
H4: the policy support has significant positive influence on port transformation and upgrading. That is, the greater the support of government policies in the transformation and upgrading, the better economic, social and environmental benefits brought by the transformation and upgrading.
H5: the financial support has significant positive influence on port transformation and upgrading. That is, the greater the support of financial institutions for the transformation and upgrading, the better the economic, social and environmental benefits brought by the transformation and upgrading.

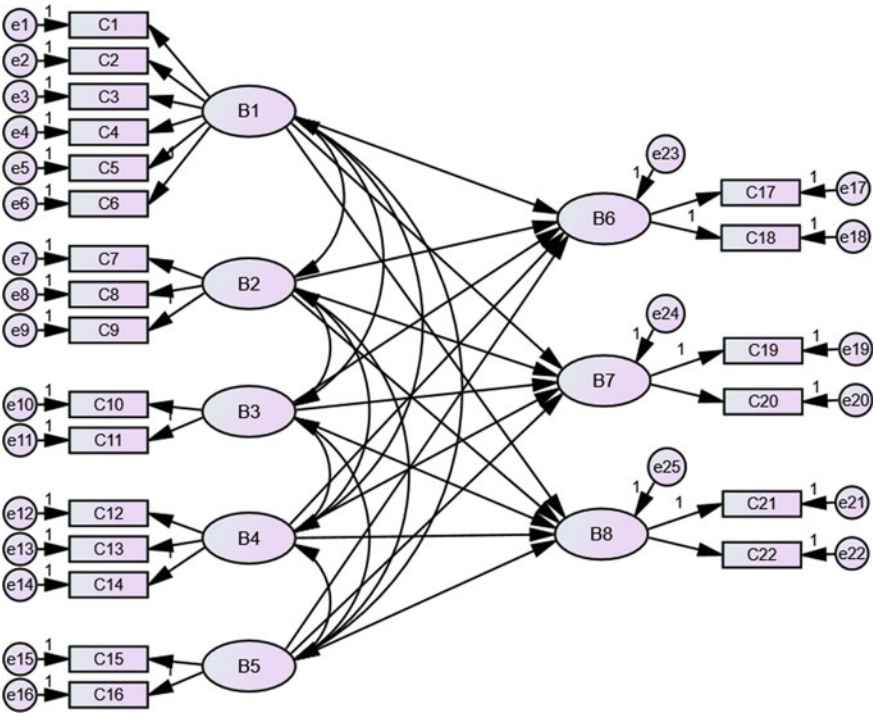


Fig. 1 Structural equation model of influence factor of port transformation and upgrading

4.2 Model Establishment

Based on the influence factors and model assumptions selected by the previous study, the initial structural equation model is constructed as shown in Fig. 1. In Fig. 1, the ellipse represents the potential variable, the rectangle represents the observation variable, and the arrow represents the path between variables.

4.3 Data Processing

The data were collected by questionnaire. According to the previous study, 22 items were set up in the questionnaire. The 7 level scale, namely 'very non-impact/conforming', 'no impact/conforming', 'less impact/conforming', 'general', 'comparative impact/conforming', 'impact/conforming', 'very impact/conforming', were given value 1, 2, 3, 4, 5, 6 and 7 respectively, to reflect the influence degree of the 5 factors mentioned above to the port transformation. To ensure the awareness of the investigators on the knowledge of port transformation and upgrading, the survey

samples were collected from the port industry managers, port professional consultants, port enterprise managers and the related users, such as shipping and freight forwarding. In this survey, 300 paper and electronic questionnaires were distributed, among which 56 of the incomplete information questionnaires were discarded, and 244 valid questionnaires were obtained. The effective recovery rate was 81.33%.

4.3.1 Reliability Analysis

Reliability refers to the stability and consistency of the results measured by the questionnaire. The Cronbach's α is used to analyze the internal reliability of the questionnaire in this article. It is generally believed that when the Cronbach's α is greater than 0.7, the data is high reliable. SPSS23.0 is used to analyze the reliability of the questionnaire, and the Cronbach's α was 0.764, greater than 0.7, indicating that the questionnaire reliability is good, and the consistency and reliability of the sample data are acceptable.

4.3.2 Validity Analysis

Validity refers to the extent to which the measured results of a questionnaire can reflect the desired content. Factor analysis is usually used to test the validity of the measure index of each latent variable for it can determine whether there is a strong correlation between each measure index of the same latent variable. Software SPSS 23.0 is used to test the validity of the questionnaire, it shows that the KMO is 0.715, and the significance probability of Bartlett's Test is 0, indicating that the validity of the designed questionnaire was accepted.

4.4 Model Solution

The software AMOS 24 is used to analyze the structural equation model and data. The fitting index, as shown in Table 2, is obtained through the modified model.

It can be found in Table 2 that each fitting index of the established model is ideal, and the overall fitting degree of the model is good, which shows that the model can be accepted and can effectively reflect the relationship between the influencing factors of the port transformation and upgrading.

Table 2 Model fitting index of influencing factors of port transformation and upgrading

Fitting index	GFI	RMR	RMSEA	IFI	CFI	χ^2/df	PNFI	PGFI
Standard	>0.90	<0.05	<0.05	>0.90	>0.90	1–3	>0.50	>0.50
Fitted value	0.933	0.047	0.039	0.925	0.916	1.516	0.581	0.675

The path test results of various factors affecting the transformation and upgrading of ports are listed in Table 3.

It can be seen from Table 3 that the significant level of the t test of the existing port development level and the support of government policies to the port transformation and upgrading is below 0.001. That is, the hypothesis of H1 and H4 are accepted, indicating that the existing development level of the port and the government policy support have a significant positive impact on the port transformation and upgrading. It means the higher the existing basic level of the port and the greater the support of government policies, the better the economic, social and environmental benefits brought about by the transformation and upgrading of ports. What's more, the average path coefficient of three indicators of the current development level of the port is 0.460, and the mean value of the path coefficient of the support of government policies to the evaluation of the transformation and upgrading effect is 0.638, indicating that the support of government policies has a greater impact on the port transformation and upgrading than the existing development level of the port.

The test of the three factors management of ports, management environment of ports and support of the financial organization only to economic and social benefits are passed, but it is not passed to environmental benefits. That is, the hypothesis of H2, H3 and H5 is verified by the hypothesis, that the higher the management level of the port, the better the management environment of the port and the greater the support of the financial organization, the better the economic and social benefits brought about by the transformation and upgrading of ports. Furthermore, the mean path coefficients of the management of ports, the management environment of ports and the support of financial organization to the two indexes which are the economic and social benefit of the transformation and upgrading are 0.345, 0.141 and 0.520, respectively. It indicates that the support of the financial organization has the greatest impact on the economic and social benefits, and then is management of ports, and the last is management environment of ports. The reasons for the failure of test of three factors to the environmental benefits may be due to the fact that most of the respondents consider the mentioned three factors have little impact on the environmental benefits of the port transformation and upgrading.

At the same time, the effect of each index on the factors of port transformation and upgrading is also shown in Table 4. As we can see, among the influencing factors of port transformation and upgrading, the t-test significance level of each observation index for all factors is below 0.001, means the impacts are all significant. Among the existing development level of ports, the port technology development level has the largest factor loading (0.884), followed by the port service level (0.825), the port capital strength (0.780), the port collecting and distributing transport condition (0.776), the port infrastructure (0.616), and finally, the logistics service capabilities (0.588). According to the operation and management factors, it can be found that the impact of port management level (0.782) on port transformation and upgrading is far greater than the extension of port industry chain (0.564) and diversification of port operation (0.503). Among the observed indicator of the management environment of ports, the economic situation in port hinterland (0.727) has a larger load factor than the market environment of the port and shipping industry (0.681). In the indicators

Table 3 Path test results of port transformation and upgrading

Hypothesis	Influence path	Standardized path coefficient	t	P	Conclusion
H1	Existing development level of ports → Economic benefit	0.473	6.439	***	Accept
	Existing development level of ports → Social benefit	0.459	4.421	***	Accept
	Existing development level of ports → Environmental benefit	0.449	4.123	***	Accept
H2	Management of ports → Economic benefit	0.322	3.146	***	Accept
	Management of ports → Social benefit	0.367	3.348	***	Accept
	Management of ports → Environmental benefit	0.082	1.623	0.087	Reject
H3	Management environment of ports → Economic benefit	0.160	2.219	***	Accept
	Management environment of ports → Social benefit	0.121	2.124	***	Accept
	Management environment of ports → Environmental benefit	0.056	0.229	0.268	Reject
H4	Support of government policies → Economic benefit	0.664	9.007	***	Accept
	Support of government policies → Social benefit	0.636	7.313	***	Accept
	Support of government policies → Environmental benefit	0.614	7.774	***	Accept
H5	Support of the financial organization → Economic benefit	0.535	6.532	***	Accept
	Support of the financial organization → Social benefit	0.505	5.820	***	Accept

(continued)

Table 3 (continued)

Hypothesis	Influence path	Standardized path coefficient	t	P	Conclusion
	Support of the financial organization → Environmental benefit	0.078	1.310	0.756	Reject

Note The '***' indicates that the significance level is less than 0.001

Table 4 Effect of each index on the factors of port transformation and upgrading

Factor	Observed indicator	Standardized factor loading	t	P
Existing development level of ports	Port infrastructure	0.616	–	***
	Logistics service capacity	0.588	2.173	***
	Port collecting and distributing transport condition	0.776	4.383	***
	Port service level	0.825	5.004	***
	Port technology development level	0.884	9.901	***
	Port capital strength	0.780	4.747	***
Management of ports	Port management level	0.782	4.977	***
	Diversification of port operation	0.503	2.853	***
	Extension of the port industry chain	0.564	–	***
Management environment of ports	Market environment of port and shipping industry	0.681	3.672	***
	Economic situation in port hinterland	0.727	–	***
Support of government policies	Fiscal and tax incentives	0.732	4.501	***
	Special support policy	0.733	4.732	***
	Port industry talent introduction policy	0.904	–	***
Support of the financial organization	Financial convenience	0.670	4.531	***
	Strength of bank loans	0.648	–	***

(continued)

Table 4 (continued)

Factor	Observed indicator	Standardized factor loading	t	P
Economic benefit	Port output	0.764	–	***
	Port throughput	0.408	1.672	0.008
Social benefit	Solve the employment	0.568	–	***
	Contribution to hinterland economy	0.632	3.946	***
Environmental benefit	Resource utilization	0.779	–	***
	Pollutant generation and discharge	0.639	4.305	***

Note The '***' indicates that the significance level is less than 0.001

of support of government policies, the results show that the port industry talent introduction policy is the most important (0.904), the fiscal and tax incentives (0.732) and the special support policy (0.733) are the next. And about the support of the financial organization factors, the degree of influence of financing facilities (0.670) and the strength of bank loans (0.648) on the transformation and upgrading of ports is not significantly different.

While the port throughput as an observation indicator of economic benefits does not pass the t-test. The reason may be that the survey object does not think that the port transformation and upgrading has a major impact on the port throughput improvement. All the other observed indicators pass the test, and the significance are below 0.001 all, therefore, the port throughput is no longer considered as an evaluation index of the effect of port transformation and upgrading effect. We can say that in terms of economic benefits of the effect of port transformation and upgrading is mainly reflected in the improvement of port output (0.764); in terms of social benefits, the increase in economic contribution of hinterland (0.632) is better than that of solving employment (0.568); while the environmental benefits are mainly reflected in the increase in resource utilization (0.779), followed by the decrease in pollutant generation and discharge (0.639).

In summary, according to the impact of port transformation and upgrading factors, they are ranked in order of the support of government policies, the support of the financial organization, the existing development level of ports, the management of ports, and the management environment of ports. Therefore, it can be considered that the key influencing factors of Chinese port transformation and upgrading are the support of government policies and the support of the financial organization. What's more, the influence of the existing development level of ports and the management of ports cannot be ignored, while the management environment of ports does not have a significant impact on the transformation and upgrading of Chinese ports.

5 Conclusions

In this article, the structural equation model of the influencing factors of the port transformation and upgrading is established, and the key influencing factors as well as their influence modes are identified by the obtained questionnaire survey data. The model results show that the government policy support has the most significant impact on the economic benefits of port transformation and upgrading, just as the increase of port output. What's more, it is the most important for the government to introduce port industry talent introduction policies for the port transformation and upgrading. Then, the support of financial organizations' financing and loan can provide a solid financial service guarantee for the transformation and upgrading of ports, which has an important impact on the increase of port output and its contribution to the hinterland economy. As we know, the existing development level of the port is the decisive factor for the capacity and development space. As the previous study confirmed, the technical development level, the service level, the fund strength, collecting and distributing transport condition and the construction of the port infrastructure are all important parameters, which affect the port's output, its contribution to the hinterland economy and the resource utilization by the transformation and upgrading of the port. In terms of the port management, the level of port management is the primary factor affecting the transformation and upgrading of ports. At the same time, the diversification of the port management and the extension of the port industry chain also have an impact on the transformation and upgrading of ports.

In conclusion, in order to effectively enhance the port transformation and upgrading capacity, the government departments should play its important role, that is also an important embodiment of Chinese characteristics. The port enterprises should also actively seek the support of the financial organizations and then find out the direction of the upgrading to promote the transformation and upgrading of the port, so that maximize the benefits. At last, the path analysis of port transformation and upgrading can be the direction of further research in the future.

Acknowledgements This research was supported by the National Natural Science Foundation of China (No. 41401120) and the Jiangsu Social Science Fund (No. 14JD014).

References

1. Langen PWD, Visser EJ (2005) Collective action regimes in seaport clusters: the case of the lower Mississippi port cluster. J Transp Geogr 13(2):173–186
2. Mccarthy J (1998) Reconstruction, regeneration and re-imaging: the case of rotterdam. Cities 15(5):337–344
3. Kim M, Sachish A (1986) The structure of production, technical change and productivity in a port. J Ind Econ 35(2):209–223

4. Wang WY (2017) Research on the development of the Ningbo free trade zone and the free trade zone in 'The Belt and Road Initiative' background. Doctoral dissertation, Chinese Graduate School of Social Sciences
5. Li ZW (2016) Research on the development path of Beijing Tianjin-Hebei Port Group from the perspective of "The Belt and Road". Hebei Acad J 3:139–144
6. Zhang GH (2015) The road of port transformation and upgrading under 'The Belt and Road' strategy. China Natl Cond Strength 3:17–19
7. Liang ZN (2015) Research on the transformation and upgrading development strategy of Dalian port. Doctoral dissertation, Dalian University of Technology
8. Huang LM (2014) Qingdao port transformation development strategy research. Doctoral dissertation, Qingdao Technological University
9. Zeng JW (2013) Research on the transformation and upgrading and impact factors of Hong Kong International Shipping Center. Doctoral dissertation, Jinan University
10. Hou JT (2004) Structural equation modeling and its application. Education Science Press
11. Wu ML (2009) Structural equation model: operation and application of AMOS. Chongqing University Press, China
12. Krishnakumar J, Ballon P (2008) Estimating basic capabilities: a structural equation model applied to Bolivia. World Dev 36(6):992–1010
13. Noyan F, Simsek GG (2011) Structural determinants of customer satisfaction in loyalty models: Turkish retail supermarkets. Procedia Soc Behav Sci 30:2134–2138
14. Qi L, Wei T (2013) The performance evaluation model of corporate social responsibility based on iso 26000 standard. Sci Res Manag 3:84–92
15. Feng XJ, Jiang LP, Yan YX, Cheng-Cai LI (2011) Study on core competitiveness of port group. J Transp Syst Eng Inf Technol 11(1):135–141
16. Wu G, Wang Y (2013) Influencing factors of port efficiency and evaluating indices system. Port Eng Technol 50(1):6–8
17. Huang JY, Yan YX (2004) Design scheme of comprehensive evaluation index system of port container transport competitiveness. Shipp Manag 26(9):8–10
18. Jiang LP, Feng XJ, Wang W (2011) 'Port-industry-city' composite system coordination model. J Econ Water Resour 29(1):11–14
19. Chen SX, Dai MH (2006) Ports competition evaluation model and application in Northeast Asia ports. J Dalian Marit Univ (Soc Sci Ed) 5(4):39–43
20. Guo J (2003) Research on integrated elements of port development. China Transp Rev 3:12–14
21. Lu CY (2003) Construction and application of the evaluation model of port competitiveness. Doctoral dissertation, Shanghai Maritime University
22. Liu JP, Wang C, Li KY (2015) Research on competitiveness of port in port group based on low-carbon development. Value Eng 30:77–79

Study on the Status Quo and Development of Rural Highway Traffic in South Jiangxi

Qing-Zhang Yuan, Liang-Song Zhi and Wang Jian

Abstract In recent years, with the strategy of revitalizing and developing the country's south Jiangxi region, the rural transportation in southern Jiangxi has been greatly improved, but it is still in extensive development. The rural highway traffic in south Jiangxi is not optimistic. This paper analyzes the status quo of rural highway traffic in southern Jiangxi, and comprehensively analyzes the characteristics of rural highway traffic in southern Jiangxi from the aspects of highway grade, vehicle problems, terrain and planning, and puts forward two countermeasures for local conditions and talent reserve from the perspective of long-term development.

Keywords Rural road · Traffic · Talent reserve

In March 2018, the first session of the 13th National People's Congress decided to set up the Ministry of Agriculture and Rural Development, not to retain the Ministry of Agriculture, and to raise rural construction and development to the core of national development. The Gannan region is mainly composed of 3 districts, 14 counties, and 1 county-level city under the jurisdiction of the prefecture level. The area in the southern part of Jiangxi Province accounts for about one-fourth of the total area of Jiangxi Province, and the population accounts for about one-fifth of the total population in Jiangxi Province. In June 2012, the State Council formally promulgated the State Council's "Supporting the Revitalization and Development of the Original Central Soviet Area in Southern Jiangxi and Other Regions". "Several Opinions" strive to promote all-round revitalization of the former Central Soviet Area in southern Jiangxi, including in all aspects and in all fields. So far, the revitalization and development of the Gannan region has become a national strategy, and the strategic position of Gannan in the country has become more prominent [1]. While Gannan is a less-developed area, the urbanization rate is not high, and the rural area occupies a large proportion. Therefore, rural construction and development is the key to the realization of comprehensive rejuvenation in Gannan. The data provided by the Traffic Management Bureau of the Ministry of Public Security today shows that

Q.-Z. Yuan (✉) · L.-S. Zhi · W. Jian
Jiangxi College of Applied Technology, Ganzhou 341000, China
e-mail: 2079304187@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_30

On average, there are 1690 traffic accidents per day, 257 deaths, 1147 injuries, and direct economic losses of 7.31 million yuan. According to the statistics of the Ministry of Public Security, traffic accidents on rural roads nationwide account for 30% of the total number [2]. It can be seen that rural transportation is an important part of transportation and its construction and development are particularly important.

This article starts with the problems existing in rural highway traffic in southern Jiangxi, analyzes and diagnoses the causes, and accordingly proposes the direction of traffic development.

1 Status Quo of Rural Highway Traffic in South Jiangxi

As of February 28, 2017, by the end of the year, the total population of households in Gannan was 974.25 million, and the rural population accounted for 58% of the total population, with 5.6606 million people. The total length of rural roads in southern Jiangxi is 25,832 km, accounting for 82% of the total mileage in southern Jiangxi Province, ranking first in the province. Rural roads are the main route for rural people to travel, and the development of rural roads that are closely related to them is particularly important [3].

1.1 Poor Technical Grades of Rural Roads, Incomplete Traffic Safety Facilities

Due to the constraints of economic development, although the total mileage of rural roads in the southern Jiangxi Province ranks first in the province, the technical grade of rural roads is relatively low, and the total mileage of medium-grade roads is less than 50%. Most rural roads belong to foreign highways. Some of the highways outside the country are self-financed in rural areas and lack of technology and supervision, resulting in narrow road widths, many entrances and exits, and irregular road lines (Fig. 1).

In recent years, rural road construction has been stepped up at the national level, but there is a lack of systematic planning. First, it lacks features and it is stereotyped. Second, there are no basic traffic signs.

1.2 Increase in the Number of Private Cars in Rural Areas and Weak Awareness of Drivers' Traffic Regulations

In recent years, the living standards of the rural population have risen and compared with each other, leading to a sharp increase in private cars in southern Jiangxi. As of the end of 2017, the number of private cars in the city reached 749,100, but due

Fig. 1 Typical road in Longnan district

to the low level of car owners in rural areas, the awareness of traffic safety is weak. Frequent accidents, according to statistics, rural accidents account for as much as 40% of the total, and showing an upward trend.

1.3 There Are Many Mountainous Roads in Southern Jiangxi, and the Security Protection Facilities Are Weak

Although the rural roads are being vigorously developed at present, they are still in extensive development, there is no reasonable planning, and there is no necessary safety protection facilities. On February 20, 2018, a license plate for Ruijin, Jiangxi Province, was a rural passenger bus for the B44296 (Nineteen people were loaded and 31 were actually carried out). They drove from Ruijin City to Ruilin Township, Ruijin City, passing the 319 State Road, Changsha Highway, Ningdu County, Ningfang County, Gefang Village, where they turned sideways, killing 9 people on the spot. Rescue died and 20 people were injured. On the one hand, this type of accident shows that the traffic regulations in rural areas of southern Jiangxi are weak and overloaded; on the other hand, rural roads in rural areas in southern Anhui lack the necessary safety protection facilities.

Fig. 2 T junction

1.4 The Lack of Forward-Looking Planning of Rural Roads in Southern Jiangxi and Irrational Intersections

There are few mountains and plains in southern Jiangxi, and the road surface is often very narrow. The street is basically a street building, and most of the intersections are non-crossroads and T-junctions (see Fig. 2). This kind of road can't see the opposite vehicle because of the driver. There is a large blind spot of sight and it is extremely prone to traffic accidents [4]. According to statistics, nearly 60% of traffic accidents are related to rural road conditions.

2 The Development Direction of Rural Highway Traffic in South Anhui

2.1 Based on Local Conditions, Develop Rural Transportation in Accordance with the Characteristics of Gannan

General Secretary Jinping proposed the strategy of rural revitalization at the 18th National Congress of the Communist Party of China, referring rural development to an unprecedented height and developing rural areas. There is an old saying that is good to get rich and build roads. Therefore, to achieve rural revitalization, the primary solution is to Due to the rural transportation problem, due to the large number of mountains and uneven roads in the southern Jiangxi region, we combined the characteristics of the Gannan region in the development of rural highways. First, we established a reasonable transportation layout. Most of the southern Jiangxi Province

are built on mountains, so Most rural and rural highways are winding roads. Therefore, reasonable traffic safety facilities are needed to reduce the loss of traffic accidents; but the road width of rural roads must at least reach two-way two-lane roads. Previously established roads and highways for rural roads have been unable to meet the current development Demand, especially during busy hours and holidays, traffic congestion on rural roads is very serious [5].

2.2 Establish a Think Tank for Talents to Study and Plan Rural Transportation in Southern Gannan

Talent is the driving force behind the development. Dong Mingzhu, a famous entrepreneur, puts forward: "There is no talent, everything is zero." The importance of talents in the development of the country and the city can be seen. In order to develop rural transportation in the Longnan area, we must first have relevant talent pools, and we must have long-term personnel engaged in relevant research in order to have a reasonable and distinctive rural transportation that is suitable for the development of the Gannan region. At the national level, at the 18th National Congress of the People's Republic of China, and specifically for the Ministry of Agriculture and Rural Affairs, in the southern part of the province, we also need to set up relevant research institutions specifically to promote cooperation with major scientific research institutions in southern Jiangxi and establish relevant research institutions. Set up related majors, carry out talent reserve and knowledge reserve.

3 Conclusion

It is imperative to strengthen rural highway construction in the Gannan region and is of great significance for the revival of the former Central Soviet Area in southern Jiangxi. Similarly, in line with local conditions, strengthening talent pools is a solution to the construction of rural transportation in southern Jiangxi. Starting from the characteristics, setting foot in personnel training, and comprehensively adopting countermeasures to improve the status quo of rural transportation in southern Jiangxi to achieve the overall coordination of rural highways in southern Jiangxi. Continuous development.

References

1. Matthias G, Hu Y, Zhou X, Ning B, Mi J, Liu J, Wang G, Wang J, Dong C, Zhang L (2018) Road traffic safety development report (2017). China Emerg Manag 02:48–58
2. Some opinions of the State Council on supporting the rejuvenation and development of the former Central Soviet Area in Southern Jiangxi Province etc. (2012). Old District Construction (13):27–36
3. Ou J (2004) Current rural road traffic status, management defects and countermeasures. J Sichuan Police Acad 05:64–68
4. Li J (2009) Discussion on the basic method of rural highway network planning. China Eng Consult 08:18
5. Ju X, Hu B, Chen F (2002) Discussion on design method of rural highway network planning layout. J Highw Transp Res Dev 02:138–141

Analysis of Child Pedestrians' Unsafe Road Crossing Behavior at Intersections in School Zones

Lianning Fu and Nan Zou

Abstract The characteristics of child pedestrians' unsafe crossing behavior and the effects of age, gender, accompanying adults and peers on their crossing behavior were examined in this study. Three signalized and one unsignalized intersections in elementary school zones were selected to conduct the observations. Chi-square tests were used to examine the four types of child pedestrians' unsafe crossing behavior. Children in Grade 4–6 group committed more walking outside of crosswalk and running to cross the road behavior than children in Grade 1–3. Boys in Grade 4–6 were more likely to run and cross against the red light than the girls in the same age. Children accompanied with peers were more likely to walk outside of crosswalk and not look both directions before crossing. Children accompanied by grandparents were less likely to run to cross the road, but children who crossed without adults or peers committed more running behavior. Children showed more unsafe crossing behavior at signalized intersection than at unsignalized intersections. Age, gender and accompaniment have different effects on elementary school children's road crossing behavior. The insights from this study may provide reference for researchers, educators and decision-makers, to understand the characteristics of children's road crossing behavior and improve their road safety.

Keywords Children · Pedestrian · Unsafe crossing behavior · Road safety

L. Fu (✉)
School of Transportation and Logistics Engineering, Shandong Jiaotong University, 5001 Haitang Road, Jinan 250357, China
e-mail: fulianning@126.com

N. Zou
School of Control Science and Engineering, Shandong University, 17923 Jingshi Road, Jinan 250061, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_31

1 Introduction

Child pedestrians are the most vulnerable group among road users, they are more likely to be involved in traffic injuries. Especially in low-income and middle-income regions, because they walk to school more often [33], between 30 and 40% of children suffering road traffic injuries are pedestrians in these countries [32]. In China, 1878 child pedestrians involved in traffic injuries for 6–11 year olds in 2013, accounting for 44.34% of the total number of children injured in this age group [19]. Therefore, the safety of child pedestrians still calls for continuous and widely concern.

Children have limitations in size, vision, hearing, attention and judgment, which make it difficult for them to make correct crossing decision in the same way as adults. Moreover, limited cognitive skills, risk-taking, as well as peer pressure, also contribute to children's vulnerability in traffic safety [10, 21, 26, 32]. Walking is the main transport mode to school for elementary school children in many countries, which makes them have a higher level of exposure in the complex traffic [7–9, 15]. Previous research identified that pedestrian collisions were more likely to occur at intersections or pedestrian crossing sites [23], and pedestrians' violation behavior was one of the most important causes for pedestrian collisions [18, 27, 30]. Agran et al. [1] analyzed the data of child pedestrians' traffic injuries aged 0–14 years and found that 81% of children were injured at midblock or intersections. Therefore, it's important to improve child pedestrians' safe road crossing behavior on the prevention of traffic injury.

Young children have poor ability to recognize danger when they are walking and crossing, and many of them are accompanied to school by an adult [2, 9, 29]. With the increase of age, they have better cognitive abilities, show more adult-like choices and allocate their attention during walking and crossing [10, 28, 29]. However, older children were more likely to take risk-taking behavior [34].

Gender differences have always been in pedestrians' behavior and traffic injury all over the world, males have significant higher traffic injury mortality rates than females and more risk-taking behavior [14, 20, 22]. Similarly, Barton and Schwebel [3] examined the roles of gender in children's pedestrian behavior, the results showed that boys were more likely to behave in a risky manner than girls. Rosenbloom et al. [25] found that girls had better performance in positive crossing behavior, such as stop, look, cross straight and cross in crosswalk. These differences were also confirmed in other studies [12, 31, 36].

In addition to age and gender differences, Rosenbloom [24] found that pedestrians were more likely to cross against the red light when they were standing alone than waiting with others, it could be explained by the theory of Social Control. However, this mechanism might not be fully applicable to children. Ben-Moshe [5] found that child pedestrians showed more risk-taking crossing behavior when they were with peers than they walked alone. Rosenbloom et al. [25] studied children's crossing behavior and found that children crossing alone performed better in positive crossing behavior than those crossing with peer-group members, such as stopping and looking before crossing, as well as crossing straight head.

In China, walking was the one of the most common school commute mode for children in elementary school [16], and many of them were escorted to school by adults [13]. Although previous studies have investigated child pedestrians' behavior, unlike many other countries, children in elementary school are more common accompanied to school by their grandparents than their parents in China. But a small number of studies give attention to these children's crossing behavior on their school commute, especially unsafe behavior. Therefore, the present study attempts to explore the characteristics of elementary school children's unsafe crossing behavior at signalized and unsignalized crosswalks and examine the differences with regard to age, gender and accompaniment.

2 Method

2.1 Participants

A total of 1084 children in elementary school were observed at four marked crosswalks in school zones, including 568 (52.4%) boys and 516 (47.6%) girls. Of all the participants, 42.8% of children were estimated as Grade 1–3 and 57.2% were estimated as Grade 4–6. Children's age was mainly estimated by their heights and the time to leave school. In China, in order to avoid the traffic jams around school zone, most of elementary schools arrange children in different grades leave school at different time. Generally, the low grades (Grade 1–3) are allowed to leave school first, and then the high grades (Grade 4–6). Moreover, 15.7% of the children were accompanied by parents, 23.4% were accompanied by grandparents, 36.5% were in a group with peers, and 24.4% walked alone.

2.2 Location

The observations were carried out in four intersections in different school zones in Jinan City. Most of intersections near school were signalized, so three signalized marked crosswalks and one unsignalized marked crosswalks were selected. The distance between the intersection and school gate ranged between 65 and 160 m.

2.3 Procedure

Each intersection was observed for two periods of a day, from 11:00 a.m. to 12:00 noon and 4:00 p.m. to 5:00 p.m., hours when children were on the way from school to home. The observations were conducted on weekdays, it took eight days to collect

the data. The weather was warm and sunny during the observation sessions. Each intersection was arranged with two trained investigators, using cameras to record children's crossing behavior. They were trained to collect data based on the same criteria, including classify the children's gender, age, types of combination, as well as unsafe crossing behavior. There were four types of children combination: (a) child accompanied by parents, (b) child accompanied by grandparents, (c) child accompanied with peers, and (d) child alone. Besides, four types of unsafe crossing behavior were observed and recorded: (a) walking outside of the crosswalk, (b) not looking before crossing, (c) running to cross the road, and (d) crossing against the red light.

3 Results

The four types of child pedestrians' unsafe crossing behavior were analyzed and scored in this study. Children who had one type of unsafe crossing behavior would receive a score of 1. The frequency and percentage of each type of unsafe crossing behavior were presented in Table 1. The results indicated that not looking before crossing and not walking inside the crosswalk were the two most frequent unsafe crossing behavior, followed by the running to cross and crossing against the red light. We calculated three types of unsafe crossing (i.e., walking outside of crosswalk, not looking before crossing and running to cross the road) among all the children, the mean score for the three behavior was 1.44 (S.D. = 0.76). The results revealed that only 7.8% of children scored 3, namely did not have any unsafe crossing behavior. Worse still, 8.6% of them committed all the three unsafe crossing behavior. Besides, 47.1% of them had one unsafe crossing behavior and 36.4% had two unsafe crossing behavior.

Table 1 The frequency and percentage of different types of child pedestrians' unsafe crossing behavior

Unsafe crossing behavior	Frequency	Percentage
Walking outside of crosswalk (n = 1084)	420	38.7
Not looking before crossing (n = 1084)	876	80.8
Running to cross the road (n = 1084)	260	24.0
Crossing against the red light (n = 732)	85	11.6

Table 2 Age differences in child pedestrians' unsafe crossing behavior (%)

	Grad 1–3	Grad 4–6	χ^2-value
Walking outside of crosswalk	151 (32.5)	269 (43.4)	$\chi^2(1) = 13.149^{***}$
Not looking before crossing	422 (90.9)	454 (73.2)	$\chi^2(1) = 53.755^{***}$
Running to cross the road	76 (16.4)	184 (29.7)	$\chi^2(1) = 25.741^{***}$
Crossing against the red light	29 (12.3)	56 (11.3)	$\chi^2(1) = 0.155$

$^{***}p < 0.001$

3.1 Age Differences

The age of children was estimated by the methods mentioned above, Chi-square tests were conducted to determine whether there were significant differences in unsafe crossing behavior between the two age groups. The results in Table 2 showed that three unsafe crossing behavior were found differences between Grade 1–3 group and Grade 4–6 group, walking outside of crosswalk, not looking before crossing and running to cross the road. Children in Grade 4–6 group were more likely to walk outside of crosswalk (χ^2 $(1, N = 1084) = 13.149$, $p < 0.001$) and run to cross the road (χ^2 $(1, N = 1084) = 25.747$, $p < 0.001$). However, children in Grade 1–3 group were more likely not to look left and right before they cross the road, despite the proportions of this unsafe crossing behavior were both higher for all the children, χ^2 $(1, N = 1084) = 53.755$, $p < 0.001$. Moreover, there was no significant difference between the two age groups in the crossing against the red light, χ^2 $(1, N = 732) = 0.155$, $p > 0.05$.

3.2 Gender Differences

The Chi-square test results showed that there were no significant differences in the four types of unsafe crossing behavior between boys and girls among Grade 1–3 group (see Table 3). One possible reason was that most of children in this age group were accompanied by their parents or grandparents, their crossing behavior dependent on the adult's behavior to a certain extent. Whether boys or girls, most of them just follow the adults when crossing the road.

It was found that 93.5% of children in Grade 4–6 group were not accompanied by adults. Therefore, children who are accompanied by adults are not analyzed in this age group. Chi-square test results in Table 3 indicated that boys had more unsafe crossing behavior than girls. Boys were more likely to run during crossing the road than girls, χ^2 $(1, N = 580) = 6.502$, $p < 0.05$. Besides, more boys crossed against the red light than girls at signalized crosswalks, χ^2 $(1, N = 465) = 8.693$, $p < 0.01$.

Table 3 Gender differences in child pedestrians' unsafe crossing behavior (%)

	Grade 1–3			Grade 4–6		
	Boys	Girls	χ^2-value	Boys	Girls	χ^2-value
Walking outside of crosswalk	92 (36.4)	59 (28.0)	$\chi^2(1) = 3.699$	126 (42.6)	124 (43.7)	$\chi^2(1) = 0.017$
Not looking before crossing	230 (90.9)	192 (91.0)	$\chi^2(1) = 0.001$	222 (75.0)	201 (70.8)	$\chi^2(1) = 1.311$
Running to cross the road	45 (17.8)	31 (14.7)	$\chi^2(1) = 0.804$	105 (35.5)	73 (25.7)	$\chi^2(1) = 6.502^*$
Crossing against the red light	17 (13.8)	12 (10.6)	$\chi^2(1) = 0.560$	37 (15.7)	16 (7.0)	$\chi^2(1) = 8.693^{**}$

$^*p < 0.05$, $^{**}p < 0.01$

With regard to the two other types of unsafe crossing behavior, walking outside of crosswalk and not looking before crossing, no significant differences between boys and girls were found among Grade 4–6 group.

3.3 Individual and Groups Differences

The Chi-square test results in Table 4 revealed that children crossing with their peers were more likely to walk outside of crosswalk than children crossing alone and accompanied by adults, χ^2 $(3, N = 1084) = 14.717$, $p < 0.01$. Children crossed without adults looked more before crossing than children accompanied by parents and grandparents, χ^2 $(3, N = 1084) = 77.361, p < 0.001$. More than 90% of children

Table 4 Individual and groups differences in child pedestrians' unsafe crossing behavior (%)

	Walking outside of crosswalk	Not looking before crossing	Running to cross the road	Crossing against the red light
Parent-child pair	59 (34.7)	161 (94.7)	19 (11.2)	5 (6.0)
Grandparent-child pair	86 (33.9)	234 (92.1)	23 (9.1)	21 (14.2)
Child peer group	183 (46.2)	303 (76.5)	110 (27.8)	29 (9.8)
Child alone	92 (34.8)	178 (67.4)	108 (40.9)	30 (14.7)
χ^2-value	$\chi^2(3) = 14.717^{**}$	$\chi^2(3) = 77.361^{***}$	$\chi^2(3) = 90.948^{***}$	$\chi^2(3) = 6.432$

$^{**}p < 0.01$, $^{***}p < 0.001$

Table 5 Signalized and unsignalized differences in child pedestrians' unsafe crossing behavior (%)

	Walking outside of crosswalk	Not looking before crossing	Running to cross the road
Signalized crosswalk	327 (44.7)	616 (84.2)	204 (27.9)
Unsignalized crosswalk	93 (26.4)	260 (73.9)	56 (15.9)
χ^2-value	$\chi^2(1) = 33.363^{***}$	$\chi^2(1) = 16.229^{***}$	$\chi^2(1) = 18.648^{***}$

$^{***}p < 0.001$

accompanied by adults did not look before crossing, however, we found that only 38.8% of parents and 36.2% of grandparents took a look before crossing when they crossing with their children. Moreover, children crossed with peers were less likely to look left and right before crossing than children walked alone. Compared to children accompanied by adults and peers, children crossing alone committed more running behavior, on the contrary, children accompanied by grandparents were less likely to run to crossing, $\chi^2\,(3, N = 1084) = 90.948, p < 0.001$. Furthermore, children walked alone and accompanied by grandparents performed worse in crossing against the red light, but the differences were not significant at the 0.05 level ($\chi^2\,(3, N = 732) = 6.432, p > 0.05$).

3.4 Signalized and Unsignalized Differences

Children's unsafe crossing behavior at signalized and unsignalized crosswalk were also examined using Chi-square test. The results in Table 5 revealed that children showed more unsafe crossing behavior at signalized crosswalk compared to unsignalized crosswalk, including walking outside of crosswalk ($\chi^2\,(1, N = 1084) = 33.363$, $p < 0.001$), not looking before crossing ($\chi^2\,(1, N = 1084) = 16.229, p < 0.001$), and running to cross the road ($\chi^2(1, N = 1084) = 18.648, p < 0.001$).

4 Discussion

The present study observed elementary school children's road crossing behavior on their school commute way. The findings revealed that more than 90% of younger children (Grade 1–3) were not looking left and right before crossing the road. This may be in part due to they are lack of cognitive abilities to the traffic risk and not formed good crossing habits. Zeedyk et al. [35] conducted a study to observe the behavior of children in Grade 2 of elementary school at two different road crossings, the results showed that no more than 41% of the children looked for the traffic,

which was considered to be poor performance. Unfortunately, the proportion of this crossing behavior in our study was much lower than their results. One possible reason is that most of children in Grade 1–3 are accompanied by parents and grandparents, they may not have been educated to look for the traffic before crossing. Children in Grade 4–6 were more likely to look left and right than children in Grade 1–3. This consistent with the study by Dunbar et al. [10]. They found that the older children were able to switch attention more rapidly and more likely to look at traffic when they crossing. However, in this study we found that the older children in Grade 4–6 had more risky crossing behavior (i.e. running to cross) than the younger children in Grade 1–3.

Gender has been identified as one of the important factors on children's traffic behavior and injuries [3, 11, 25, 37]. The results of present study showed that the gender differences only exists in the older children (Grade 4–6). Not surprisingly, boys committed more unsafe crossing behavior than girls, which was consistent with previous studies.

In addition, most of children observed in Grade 1–3 were accompanied by their parents and grandparents. Thus, the parents and grandparents would affect their children's crossing behavior. Children who accompanied by parents had more unsafe crossing behavior than those accompanied by grandparents. For example, children accompanied by grandparents were less likely to run to cross. Perhaps the biggest reason is that the grandparents themselves walk relative slowly, it's hard for them to cross rapid when they take their children. Also differences in children's unsafe crossing behavior were found between children accompanied with their peers and children crossed alone. Children crossed with peers were more likely to walk outside of crosswalk and not look before crossing, compared to children crossed alone. From the observations, we found that many of children talked to each other when they walked with peers in a group. This might be distracting their attention from the traffic safety. In other previous studies, children were found to take more risky behavior when they were with peers [6, 17, 32]. Nevertheless, we found more children ran to cross the road when they were alone. The proportion running behavior of children without accompaniment was significant higher than other children. One possible explanation is that there is no talking to distract their attentions when they cross alone, meanwhile, children are by nature lively and energetic, so more running behavior occur when they cross the road alone.

Based on these findings, the crossing behavior of parents, grandparents and all the children are suggested to be improved. Younger children show poor crossing skills, therefore campaigns, educations and trainings should develop to improve their crossing abilities. Thomson et al. [29] carried out the training experiences and the results showed that children's cognitive abilities to danger on the road were significant improved after the trainings. Similarly, Barton et al. [4] also found that child pedestrian would behave more safely by the simple skill training. In fact, parents and grandparents have more opportunities to teach their children in safe crossing behavior. However, they have responsibility to be a good role model for their children and then teach the children. Unfortunately, such teaching behavior was not observed in

this study. Moreover, children committed more unsafe crossing behavior at intersections, which showed their weak safe consciousness to a certain extent. At signalized intersection, there also exist potential collision risk. So children should not relax and ignore the traffic rules.

Lastly, the present study has several limitations. Studying children's crossing behavior by observation method is one of the limitations. Although this method is commonly used, the observed data may be affected by observers' bias, such as age. Besides, it may lose some other demographic information, as well as the attitude and behavior intention. Therefore, questionnaire survey should be carried out in future research, combining with observations to obtain data systematically. In addition, more samples are needed in future study, especially the data of children's crossing behavior at unsignalized intersections in school zones. The difference in children's crossing safety between cities should also be paid attention to.

References

1. Agran PF, Winn DG, Anderson CL (1994) Differences in child pedestrian injury events by location. Pediatrics 93(2):284–288
2. Ampofo-Boateng K, Thomson JA, Grieve R, Pitcainr T, Lee DN, Demetre JD (1993) A developmental and training study of children's ability to find safe routes to cross the road. Br J Dev Psychol 11(1):31–45
3. Barton BK, Schwebel DC (2007) The roles of age, gender, inhibitory control, and parental supervision in children' s pedestrian safety. J Pediatr Psychol 32(5):517–526
4. Barton BK, Schwebel DC, Morrongiello BA (2007) Brief report: increasing children's safe pedestrian behaviors through simple skills training. J Pediatr Psychol 32(4):475–480
5. Ben-Moshe D (2004) Detection of danger signals on the road and social facilitation. Bar-Ilan University
6. Christensen S, Morrongiello BA (1997) The influence of peers on children's judgments about engaging in behaviors that threaten their safety. J Appl Dev Psychol 18(4):547–562
7. Costa FF, Silva KS, Schmoelz CP, Campos VC, de Assis MAA (2012) Longitudinal and cross-sectional changes in active commuting to school among Brazilian schoolchildren. Prev Med 55(3):212–214
8. Cui Z, Bauman A, Dibley MJ (2011) Temporal trends and correlates of passive commuting to and from school in children from 9 provinces in China. Prev Med 52(6):423–427
9. Department for Transport (2015) National Travel Survey 2014: Travel to school. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/476635/travel-to-school.pdf
10. Dunbar G, Hill R, Lewis V (2001) Children's attentional skills and road behavior. J Exp Psychol Appl 7(3):227–234
11. Granié MA (2007) Gender differences in preschool children's declared and behavioral compliance with pedestrian rules. Transp Res Part F: Traffic Psychol Behav 10(5):371–382
12. Granié MA (2009) Effects of gender, sex-stereotype conformity, age and internalization on risk-taking among adolescent pedestrians. Saf Sci 47(9):1277–1283
13. Han J, Guo GZ, Li HQ (2011) Characteristics and management strategies for commuting trips of primary school students. Urban Transp China 9(2):74–79
14. Kingma J (1994) Age and gender distributions of pedestrian accidents across the life-span. Percept Mot Skills 79(3):1680–1682

15. Koekemoer K, Van Gesselleen M, Van Niekerk A, Govender R, Van As AB (2017) Child pedestrian safety knowledge, behaviour and road injury in Cape Town, South Africa. Accid Anal Prev 99:202–209
16. Liu A, Hu X, Li Y, Wang J, Zhao W, Zhai F, Ma G (2009) Modes of transportation to and from school in Chinese elementary and secondary students. Chin J Health Edu 25(1):8–10
17. Miller DC, Byrnes JP (1997) The role of contextual and personal factors in children's risk taking. Dev Psychol 33(5):814–823
18. Ministry of Public Security (2012) Annual statistical report of road traffic accidents in China, 2011
19. Ministry of Public Security & Chinese Centre for Disease Control and Prevention (2014) Research report on children's road traffic injury in China. People's Medical Publishing House, Beijing
20. Peden M, Scurfield R, Sleet D, Mohan D, Hyder AA, Jarawan E, Mathers C (2004) World report on road traffic injury prevention. Geneva. Available at: http://www.who.int/violence_injury_prevention/publications/road_traffic/world_report/en/
21. Percer J (2009) Pedestrian safety education: applying learning and developmental theories to develop safe street-crossing behaviors. Washington, DC
22. Razzaghi A, Zolala F (2015) Exploring the pedestrian' s behaviors in crossing the street based on gender. Austin J Emerg CritAl Care Med 2(1):1011
23. Ren F, Liu X, Rong J (2008) Traffic Engineering. China Communications Press, Beijing
24. Rosenbloom T (2009) Crossing at a red light: behaviour of individuals and groups. Transp Res Part F: Traffic Psychol Behav 12(5):389–394
25. Rosenbloom T, Sapir-Lavid Y, Hadari-Carmi O (2009) Social norms of accompanied young children and observed crossing behaviors. J Saf Res 40(1):33–39
26. Schieber RA, Thompson NJ (1996) Developmental risk factors for childhood pedestrian injuries. Inj Prev: J Int Soc Child Adolesc Inj Prev 2(3):228–236
27. Stutts J, Hunter W, Pein W (1996) Pedestrian crash types: 1990s update. Transp Res Rec J Transp Res Board 1538(1):68–74
28. Thomson JA (1991) Children's perception of safety and danger on the road. Br J Psychol 82:487–505
29. Thomson JA, Ampofo-Boateng K, Lee DN, Grieve R, Pitcairn TK, Demetre JD (1998) The effectiveness of parents in promoting the development of road crossing skills in young children. Br J Educ Psychol 68:475–491
30. Ulfarsson GF, Kim S, Booth KM (2010) Analyzing fault in pedestrian-motor vehicle crashes in North Carolina. Accid Anal Prev 42(6):1805–1813
31. Wang X, Wang L, Tremont JP (2013) Analysis of knowledge of crossing rules, self-reported behavior, and observed behavior at intersections. In transportation research board meeting. Retrieved from http://trid.trb.org/view/1241477
32. World Health Organization (2007) Youth and road safety. Geneva. Available at: http://www.who.int/violence_injury_prevention/publications/road_traffic/youth_roadsafety/en/
33. World Health Organization (2015) Global status report on road safety. World Health Organization. Geneva. Available at: http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/
34. World Health Organization, & Unicef (2008) World report on child injury prevention. Geneva. Available at: http://apps.who.int/iris/bitstream/10665/43851/1/9789241563574_eng.pdf
35. Zeedyk MS, Wallace L, Spry L (2002) Stop, look, listen, and think?—What young children really do when crossing the road. Accid Anal Prev 34(1):43–50
36. Zhang DL, Wei DH, Li DZ, Zhang DY, Li P, Bian Y, Shu S (2013) Pedestrian crossing behavior at unsignalized mid-block crosswalks around the primary school. Procedia—Soc Behav Sci 96:442–450
37. Zhou R, Horrey WJ (2010) Predicting adolescent pedestrians' behavioral intentions to follow the masses in risky crossing situations. Transp Res Part F: Traffic Psychol Behav 13(3):153–163

Investigating Private Cars Idling Behavior in Urban Areas

Lu Xing, Jie He, Chen Zhang, Ziyang Liu and Hao Zhang

Abstract The rapid mechanization in China results in excessive adverse effects recently, such as traffic congestion and air pollution. Affected by the negative effects, an increasing number of citizens decide to use their private cars only at a certain time, which leads to the urban private cars idling (UPCI) phenomenon. In order to investigate the UPCI behavior and its influence factors, this paper, taking Nanjing city in China as a case study, conducted a detailed survey including 279 private car owners. A logistic regression model was developed to investigate the impact factors related with UPCI. The result of regression indicated that the number of children in a family was an impeding factor which caused the fewer UPCI behaviors. The smaller job-housing distances and independence on vehicles, however, aggravated the UPCI phenomenon. The results of this study are beneficial to understand the UPCI behavior, and provide useful information for the effective urban transportation demand management (TDM) and necessary guidance for urban private car purchase and usage.

Keywords Urban private car · Idle · Travel behavior · Transportation demand management

L. Xing · J. He (✉) · C. Zhang · Z. Liu · H. Zhang
School of Transportation, Southeast University, Si Pai Lou #2, Nanjing 210096, People's Republic of China
e-mail: hejie@seu.edu.cn

L. Xing
e-mail: luxing@seu.edu.cn

C. Zhang
e-mail: reggie_cheung@hotmail.com

Z. Liu
e-mail: 230169550@seu.edu.cn

H. Zhang
e-mail: andyhao@seu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_32

As one of the major emerging country, Chinese national economy and urbanization develops quickly in recent years, leading to a big leap for its urban traffic mechanization. According to data from Chinese government, the number of private vehicles increases by 7 times in the past ten years, and the car parc per thousand people exceeded 125 in 2015 [1]. In addition, the number of vehicles exceed 2 million in 11 cities, including Beijing, Chengdu, Shenzhen and so on [2]. However, a different trend, the urban private car idling (UPCI) appears recently, i.e. a part of citizens choose to not use cars. For example, the mode ratio of private car in urban transportation has been declining since 2014 in Beijing [3], and the daily trips of per private car in Nanjing declined from 1.18 in 2004 to 0.98 in 2013.

Actually, the UPCI has already occurred in other countries, including the United States, the United Kingdom, Germany and Japan [4–6], but it doesn't draw too many attentions from researchers. Factors are not clear yet associated with the UPCI behavior. Thus, it is necessary to investigate this issue and analyze the influence factors on it. The primary objective of this paper is to investigate the UPCI behavior based on a survey data and logistic regression models.

1 Data Source and Methodology

1.1 Data Source

In order to find out the characteristics and influence factors of UPCI behavior, a questionnaire survey was conducted and private car owners in Nanjing were selected to complete the questionnaires in April, 2016. The questionnaire was sent by WeChat (via mobile phone online). The respondents were offered the possibility of obtaining lotteries in compensation for their participation. About 80% of the usable questionnaires were obtained in this way.

The main content of questionnaire contains information about the individual socio-economic properties, private car-use features, attitude to private car-use, UPIC features and so on. A total of 279 usable questionnaires and 8370 valid data were obtained after the removal of invalid questionnaire. The essential features of the investigation sample are similar to the result of residents' travel survey in Nanjing Transportation Development Annual Report [7], which indicates the sample in this study is representative.

1.2 Logistic Regression Model

Logistic regression is a mathematical statistics analysis method, which can investigate the relations between categorical dependent variable (DV) and independent

variable (IV). It includes binary and multinomial logistic regression models associated with the number of category. This study analyzes the influence factors of UPCI by developing a binary logistic regression model, as shown in Eq. (1) [8]. The DV is the uncertainty of whether to have the UPCI behavior or not. It is divided into two types according to the average days of UPCI per month (days/month). When DV equals to 0, it represents for no UPCI behavior, while equaling to 1, it represents the existing of the UPCI behavior.

$$Logit(P) = \alpha + \sum_{i} \beta_i X_i \tag{1}$$

where α is a constant, X_i is the ith IV, β_i is the regression coefficient of the ith IV (Table 1).

2 Results and Discussions

Logistic regression models are developed to evaluate the impacts of different variables on the likelihood of UPCI behavior in this study. The dependent variable (DV) is private car idling decision, and takes a value of 1 if the UPIC occurs and 0 otherwise.

It is necessary to take a preliminary screening of independent variables ahead of logistic regression analysis. Through the correlation analysis of each two factors, one of the factors must scratched out when they have strong relation. Furthermore, due to there are categorical variables among IVs, this study uses the Spearman correlation analysis function by SPSS (Statistical Package for Social Sciences) software to identify the correlation among various factors [9]. The result shows that there seems to be no discernible correlation-ship among IVs except three pairs of factors. These three pairs have moderate correlation respectively. They are 'age' and 'driving years' (correlation index = 0.64), 'year of owing private car' and 'driving years' (correlation index = 0.51), 'degree of dependence on private car-use' and 'attitude to private car-use' (correlation index = 0.50). All of their significant levels are less than 0.05, so the factor X7 (driving years) and X16 (attitude to private car-use) are eliminated on the filter condition of correlation index more than 0.5.

2.1 Univariate Logistic Regression Analysis

After completing preliminary screening, the logistic regression models can be used to analyze the remaining variables. All the variables can be analyzed simultaneously if the sample quantity is enough. But if it is not enough, it's necessary to develop univariate logistic regression analysis firstly to get more accurate results and eliminate no statistically significant variables. Then, the multivariate logistic regression analysis can be developed in the next step [10].

Table 1 The DV and IVs in the logistic regression model

Dependent variable (DV)		
Whether having the UPCI behavior, according to the average days of UPCI per month (days/month)	Categorical variable	≤10 days/month = 0
		≥10 days/month = 1

Independent variable (IV)		
Variable name	Data type	Variable-value
X1 gender	Categorical variable	Male = 1; Female = 2
X2 age	Ordered category	18–25 = 1; 26–30 = 2; 31–40 = 3; 41–50 = 4; 51–100 = 5
X3 number of adults in the family	Continuous variable	Numerical value (1–10)
X4 number of minors in the family	Continuous variable	Numerical value (0–10)
X5 number of drivers in the family	Continuous variable	Numerical value (1–10)
X6 annual household gross income (10,000 CNY/year)	Ordered category	≤15 = 1; 15–25 = 2; 25–40 = 3; 40–50 = 4; ≥50 = 5
X7 driving years (year)	Ordered category	≤3 = 1; 3–5 = 2; 5–10 = 3; 10–15 = 4; 15–20 = 5; ≥20 = 6
X8 year of owing private car (year)	Ordered category	≤3 = 1; 3–5=2; 5–10 = 3; 10–15 = 4; 15–20 = 5; ≥20 = 6
X9 number of household private cars	Continuous variable	Numerical value (1–5)
X10 number of household bikes and electric bikes	Continuous variable	Numerical value (0–10)
X11 job-housing distance (km)	Ordered category	≤2 = 1; 2–5 = 2; 5–10 = 3; 10–20 = 4; 20–30 = 5; ≥30 = 6
X12 residential area	Ordered category	Old town = 1; New town = 2 Suburb = 3
X13 distance between residence and public transport service calculated by walking time (min)	Continuous variable	Numerical value (0–60)
X14 average cost of car-use per year (10,000 CNY/year)	Ordered category	≤1.5 = 1; 1.5–2=2; 2–2.5 = 3; 2.5–3 = 4; ≥3 = 5
X15 degree of dependence on private car-use	Ordered category	Strongly dependent = 1
		Generally dependent = 2
		Dependent = 3
		Generally independent = 4
		Strongly independent = 5
X16 attitude to private car-use	Ordered category	Strongly enjoy car-use = 1
		Enjoy car-use = 2
		Use car only when needing it = 3
		Avoid car-use = 4
		Hate car-use = 5

Table 2 The result of univariate logistic regression analysis

Variable	Variable's name	Sig.	Variable	Variable's name	Sig.
X1	Gender	0.640	*X9*	Number of household private cars	0.006
X2	Age	0.706	X10	Number of household bikes and electric bikes	0.279
X3	Number of adults in the family	0.661	*X11*	Job-housing distance (km)	0.000
X4	Number of minors in the family	0.002	X12	Residential area	0.844
X5	Number of drivers in the family	0.965	X13	Distance between residence and public transport service calculated by walking time (min)	0.830
X6	Annual household gross income (10,000 CNY/year)	0.001	*X14*	Average cost of car-use per year (10,000 CNY/year)	0.001
X8	Year of owing private car (year)	0.009	*X15*	Degree of dependence on private car-use	0.000

This study conducts univariate logistic regression analysis of the IVs respectively by SPSS software. The filter of significant factors is Sig. < 0.05. However, due to the interaction among variables, the univariate analysis result may be different from the multivariate analysis result. Therefore, the filter condition should be relaxed to avoid missing important factors, so we set the filter is Sig. ≤ 0.15 in univariate logistic regression analysis. It is clear from Table 2 that the Sig. of X4, X6, X8, X9, X11, X14, X15 are less than 0.15.

2.2 Multivariate Logistic Regression Analysis

On the base of univative logistic regression analysis, the variables (X4, X6, X8, X9, X11, X14, X15) can be substituted into the multivariate logistic regression model to identify the significant influence factors. The results of stepwise regression are given in Table 3. On the screening condition of Sig. ≤ 0.05, three factors significantly affect the UPCI: X4 (Number of minors in the family), X11 (Job-housing distance) and X15 (Degree of dependence on private car-use).

The results of the examination on whether these three variables can be removed from the regression model are given in Table 3. The Sig. of three variables are all less than 0.05, it represents that they have obvious influence and cannot be removed. According to results in Table 3, the accuracy rates of idling and no idling prediction

Table 3 Result of multivariate logistic regression analysis

Variable	Coefficient (β)	S.E	Wald Chisq	df	Sig.	Odds ratio exp. (β)
X4 number of minors in the family	−0.636	0.240	7.024	1	0.008	0.530
X11 job-housing distance	−0.210	0.093	5.055	1	0.025	0.811
X15 degree of dependence on private car-use	0.636	0.118	28.966	1	0.000	1.889
Constant	−0.636	0.240	7.024	1	0.008	0.530

Model if term removed

Variable	Model log likelihood	Change in −2 log likelihood	df	Sig. of the change
X4 number of minors in the family	−168.672	7.909	1	0.005
X11 job-housing distance	−167.291	5.148	1	0.023
X15 degree of dependence on private car-use	−180.920	32.407	1	0.000

Error analysis of logistic regression model

Whether idling	0	1	Percentage correct (%)
0	101	43	70.1
1	33	102	75.6
Overall percentage (%)	–	–	72.8

in this regression model reach to 70.1% and 75.6%, respectively, and the average accuracy rate hits 72.8%, which indicates the model is credible.

It is apparent from Table 3, the Sig. of X4 is 0.008 ($\leq$0.05), which represents the number of minors in the family has significant effect on UPCI behavior. The OR of X4 is 0.530 ($\leq$1), which indicates the possibility of UPCI will decrease followed by the increase of the number of household minors. The OR of X11 is 0.811 ($\leq$1), which also indicates that the larger job-housing distance is, the fewer UPCI behaviors there will be. This conclusion can also be found in the cross analysis in the survey descriptive statistics.

Recently, many Chinese cities has the outspread urban morphology, from the central old town to new parts, such as satellite areas. However, the urban function zones are not developed reasonably. It causes the abnormal phenomenon that the employments centralize in old town while the residences disperse out of the city center. In that case, the jobs-housing distances are far generally, and it causes the heavy traffic flow goes into/out old towns at the peak hours, and severe traffic congests

consequently. The above results indicate that small jobs-housing distances can lead to the UPCI behaviors and reduce the private cars trips, which will alleviate congestion and improve traffic compositions. Thus, it is feasible to balance the job-housing distance and guide UPCI behaviors through reasonable land planning, function zones layout, comprehensive community design, and supporting facilities, thus to manage the vehicle travel demands and regularize urban transportation.

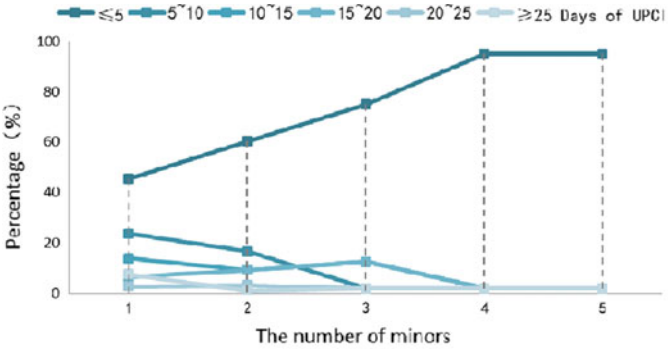
In this study, the larger value of variable X15 represents for the lower degree of dependence on private car-use (strong dependence = 1, relative dependence = 2, dependence = 3, relative independence = 4, and independence = 5), and the strong dependence condition is set as the reference. Therefore, the OR of X15 larger than 1 indicates that the lower degree of dependence on private car-use leads to more UPIC behaviors. This result is reasonable for that the degree of dependence representing for the utility preference of private car owners. And the result is also intuitive and in line with the results shown in Fig. 1.

Private cars have convenience, flexibility and other advantages. Hence, once the citizens prefer to one of the virtue of private cars, the dependence can be easily formed. In that case, they will still choose their cars in daily travels, despite of there are other alternative transportation modes. In order to balance the UPCI behaviors and reduce car travels, the public transportation systems with similar attractions can be constructed. For instance, through promoting the public transportation quality to improve the comfort and convenience, and augmenting more paths and extending service areas to improve the flexibility and accessibility of public transportation system, people will reduce the dependence on private cars and choose to idle their vehicles.

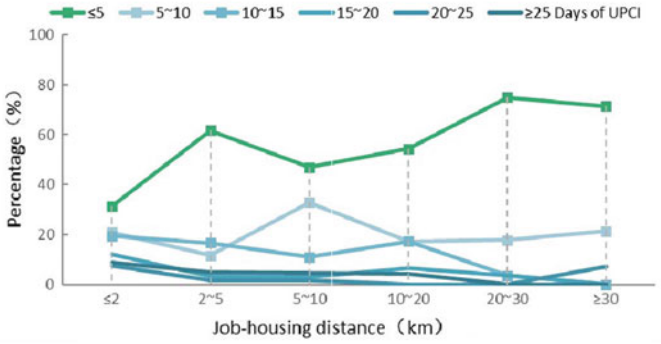
3 Conclusions

This paper, taking Nanjing city in China as a case study, conducted a detailed survey including 279 private car owners. A logistic regression model was developed to investigate the impact factors related with UPCI. The result of regression indicated that the number of children in a family was an impeding factor which caused the fewer UPCI behaviors. The smaller job-housing distances and independence on vehicles, however, aggravated the UPCI phenomenon. The results of this study are beneficial to understand the UPCI behavior, and provide useful information for the effective urban transportation demand management (TDM) and necessary guidance for urban private car purchase and usage.

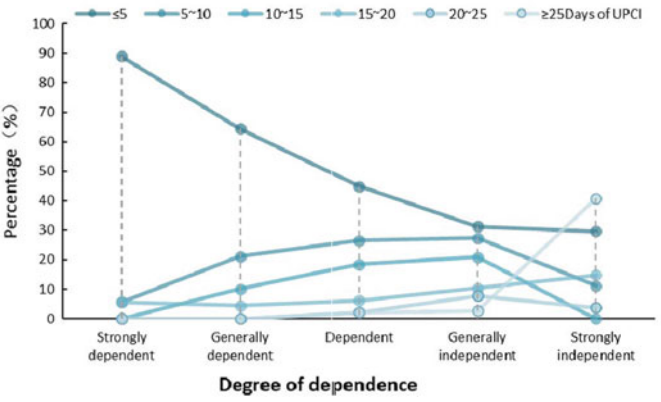
Acknowledgements The author would like to thank the Fundamental Research Funds for the Central Universities and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX17-0148). Also thanks National Natural Science Foundation of China (Grant No. 51778141 and 71601046). Their assistance is gratefully acknowledged.



（a） Number of household minors &Average days of UPCI per month



（b） Job-housing distance &Average days of UPCI per month



（c） Degree of dependence on private car-use & Average days of UPCI per month

Fig. 1 Cross analysis results of the survey data

References

1. National Bureau of Statistic of China (2016) Annual report from National Bureau of Statistics of China (2005–2015). China Statistics Press
2. National Ministry of Public Security Transportation Bureau of China (2016) 2015 National Transportation Annual Report
3. Beijing Traffic Development Research Center (2016) 2015 Beijing transportation development annual report
4. Paul A (2013) Eisenstein of the Detroit Bureau, MSNBC, Americans drive less even as economy rebounds. Accessed Sept 1 2013
5. Ye J, Chen X, Zhang H (2013) Reducing the reliance on automobiles: Portland multi-modal transportation system development. Urban Transp China 1:10–17
6. Tuttle B (2013) What happens when we reach 'Peak Car'? Time Magazine. Accessed Sep 1 2013
7. Nanjing Ministry of Public Security Transportation Bureau (2016) 2015 Nanjing transportation development annual report
8. Zhang WT, Dong W (2004) SPSS statistical analysis basics tutorial. Higher Education Press, Beijing
9. Hu Z, Liu Q (2003) Research on logistic regression and independent variables preparation in questionnaire. J Zhongnan Univ Econ Law 5:128–132
10. Wu ZQ, Wang Y, Li W (2014) Matters need attention when using logistic regression analysis. Chin Circ J 29(3):24

Study on Traffic Safety Security System at the Entrance of Middle and Primary School

Fengchun Han and Yifan Jiang

Abstract As a focal point of traffic safety, school area has received extensive attention. Especially in the peripheral areas of primary schools. During the period of schooling, motorcades, motorcycles, and non-motor vehicles are accumulating at school gates, and a large number of pedestrians are mixed there too, which cause traffic disorder. In addition, primary and secondary school students have their own characteristic of crossing the street, so the security of the elementary school area needs more targeted measures. This article starts with the analysis of road traffic characteristics and the traffic accident characteristics in front of primary and secondary schools. Combining typical real cases, this paper proposes security protection measures from traffic organization, traffic enforcement, and transportation facilities. Further constructing a traffic safety guarantee system provides an important guideline for ensuring the safety of traffic in front of primary and secondary schools.

Keywords Traffic characteristics · Traffic organizations · Pedestrian sections · Safety guarantees

1 Introduction

The transportation safety at the entrance of the school is the focus on the transportation safety management work. The continuous increase for parents to travel by vehicles, incomplete facility of the surrounding road on the campus and the incompetent enforcement management leads to the phenomena of mixed people and vehicles. Coupled with the small age of primary school students, poor safety awareness and varied course of travel, chaotic traffic orders pose a great threat to the safety of teachers and students in school. The traffic accidents of primary school students mainly occur on the way to and from school. Therefore, the significance of the

F. Han (✉) · Y. Jiang
School of Traffic Management, People's Public Security University of China, Beijing 100038, China
e-mail: hfc1966@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_33

research on the safety protection measures in front of the school is becoming more and more important.

There are not many foreign studies on the safety and security of traffic around schools. Nakitto [1] thinks that the setting more safety routes, strengthening the monitor of non-motor vehicle and road transportation safety education for the children can reduce the accident surrounding the campus. The related study of domestic, from the perspective of transportation structure and traveling time, analyzes the transportation characteristics and generally to put forward the suggestion or strengthen transportation safety from increasing transportation facility. Sun et al. [2] categorizes the middle and primary school into three types, putting forward the overall propositions to set the proper decrease facility, reasonable bus stop and facility for pedestrians, etc. Zhao et al. [3] puts forward preferred organized scheme for the traffic safety facility at the entrance of different type of school and provides the basis for traffic safety. Chang [4] studies the solutions and appraisals of traffic organization to establish the model of network appraisals and puts forward the suggestions of facility and education for the pupils.

This paper analyzes road traffic characteristics, accidental features and the factors of influencing traffic safety around the regions of middle and primary school. With the example of typical cases, it proposes the solutions of traffic safety aiming at summarizing the system of traffic safety guarantee system and has guiding value for improving the level of traffic safety around schools.

2 Traffic Characteristics at the Entrance of School

2.1 Transportation Characteristics

(1) Large quantity of transportation at the peak

The vehicles at the entrance of school are mixed and disordered. The flows of vehicles are saturated at the peak of school times.

(2) Complex traffic constitution

The pick-up vehicles is mainly based on cars. It is mixed with electric vehicles, motorcycles and bicycles. The traffic composition is complicated and the mutual interference is large.

(3) Parking problems

During the period of schooling, there is a large demand for parking around the schoolyard, and it is difficult for the existing parking spaces to meet the parking demand, resulting in significant problems in the school gates.

(4) Disordered transportation

A large number of motorized vehicles, non-motor vehicles and pedestrians gathered near the school gate during the period of schooling. Without the isolation barrier and other traffic management facilities, the phenomena of non-convergence and pedestrians crossing the street are prominent, which makes the traffic chaotic.

(5) Fast speed

The gates of urban primary and secondary schools are generally facing the trunk roads, which are in good condition with speed limits of 30 km. In the absence of a complete speed-limiting transport facility or electronic police, the speed of the vehicle will be higher than the safe speed, and the accidents can easily lead to serious consequences.

2.2 Characteristics of Transportation Participants

(1) The driving of motor vehicles is generally for middle-aged parents, who have rich driving experience and therefore drive at fast speed. They will stop at the school gate for a short time.
(2) Non-motorized drivers are generally secondary school students or parents. Parents tend to ride safely, whereas the middle school students have a strong psychological risk, such as chase racing.
(3) Pedestrians are mainly composed of primary school students and elderly parents. The seniors are poor in their ability to respond and act, and primary schools are active and have a weak sense of safety.

3 School Traffic Safety Analysis

3.1 Characteristics Analysis of Traffic Accidents

The main factors influencing the traffic safety at the entrance of middle and primary school include traffic participants, traffic management facility, illegal traffic action and environment and management, etc.

(1) Pattern of traffic accidents

According to statistic, student-related traffic accidents are mostly caused by sudden crossing of roads and disorderly walking when vehicles approaching. The accidental collision and side collision accidents caused by motor vehicles to pedestrians and non-motor vehicles are the main types of accidents.

(2) The happening time of traffic accidents

Accident occurred at noon or in the afternoon after school hours. After school, a large number of students walk or cars cross the road. The traffic order is chaotic and can easily cause accidents. Accidents often occur when students cross the street to take the bus or so.

(3) The type of traffic accidents

The accidents were concentrated in the collision of cars with electric cars, motorcycles, bicycles and pedestrians, and they suffered more injuries when riding or walking. Taking the Shenzhen in 2017 for example, there are 921 traffic accidents concerning the students in middle and primary school under the age of 15, among which 872 simple accidents, 37 injured accidents, 12 death accidents to make 9 children dead and 40 children injured. In accordance with division of traffic mode of children in the accidents, there are 13 people of taking the bus, 74 people walking, 39 people taking the motor and electric vehicles, 11 people driving the motors and electric vehicles and 21 people who take the bicycles.

3.2 The Analysis of Caused Accidents

3.2.1 The Factors of Participants in Transportation

(1) Herd mentality. This kind of psychology is irrational. Most of the pupils do not have a mature concept of traffic safety. They will only follow suit. When they see other people crossing the non-motorized lane without causing accidents or punishments, blindly following them across the road, there will be a psychological blind sense of security;
(2) Psychology is immature. Children are willing to take risks and think that they can rely on their own ability to avoid risks and get on with motor vehicles;
(3) Weak awareness of road safety. Pedestrian especially primary school students, because of their lack of knowledge of traffic safety, do not recognize the danger of crossing the motor vehicle lanes subjective.

3.2.2 Facility Factors in Transportation Management

(1) Lack of recognizing road traffic signs and markings

There are generally lack of necessary traffic sign signs around the campus. For example, the speed limit signs are not displayed repeatedly, the warning signs, prohibition signs are missing or the number of settings is insufficient. In terms of traffic marking settings, the marking lines are blurred and difficult to identify. Crosswalk lines, lane lines, center lines, stop lines, lane edge lines and traffic-induced marking lines are lacking.

(2) Non-standardized the traffic facility

The traffic signs and markings are not set according to the national standard. They are arbitrariness, lack of systematicness, continuity and standardization.

3.2.3 The Factor of Illegal Transportation

(1) Over-speed motor vehicles

Speeding is the main reason for the frequent traffic accidents in front of the school. Because the primary and middle school students are physically smaller, they are less likely to be detected by the rear vehicle when they walk through the parked vehicles, causing accidents.

(2) Motor (non-motor) vehicles parking orderless

There is limited road space at the entrance to the campus and there is a lack of necessary management measures. The non-motor vehicles are parked indiscriminately, not only occupying road space, but also infringing pedestrian rights, causing pedestrians to crowd through the middle of non-motor vehicles.

(3) Not giving-way pedestrians

At the entrance of non-light-controlled schools, vehicles in front of the school meet pedestrians or non-motor vehicles and should meet the parking rules. However, in facts, many drivers don't obey the rules; the pedestrian can only make use of the very tiny space to cross the road, which is very dangerous.

3.2.4 The Factor of Management

The traffic safety around the campus has always been the focus of traffic management work. However, the concept of traffic safety management is still relatively backward, and it is believed that chaos can bring about a lower speed and cause no casualties. During school hours, only a few traffic policemen are usually assigned to evacuate traffic in front of the campus, and there is no corresponding management of the parking of motor vehicles around the schoolyard and the phenomenon of mixed vehicles.

4 The Transportation Safety Guarantee System Construction at the Entrance of School

4.1 The Strategies of Transportation Safety

4.1.1 Optimize Transportation Facilities

(1) Pedestrian facility at the entrance of school: facilities designated crosswalk, stop line and a crosswalk matching set prompt markings.
(2) Regional speed limit and attention to children's signs: In the school road sections on both sides of the road between the opening of the 150 m area, the speed limit value should be scientific and reasonable.
(3) Mesh line: The no parking area at the entrance of the school is subject to a yellow grid marking as required. Parking is prohibited for any reason.
(4) Parking facilities: Set the parking lot for the shuttle bus, standardize parking lane marking, sign setting and parking guidance.
(5) Pedestrian overpasses or pedestrian crossings Underpasses: Pedestrian overpasses or pedestrian crossing underground passages can be set up on the road sections with large traffic flow in front of the campus.
(6) Colored road surface and reflective spikes: Use colored roads and reflective spikes to enhance visibility and induce traffic participants to regulate travel. The use of reflective spikes also improves the safety of pedestrian crossings at night.

4.1.2 Vehicle Speed Control

The use of speed limit signs in conjunction with the use of oscillating markings, visually and perceptibly reminds the driver to reduce the speed of travel and ensure that vehicles travelling on the roads around the school travel at a safe speed.

4.1.3 Traffic Separation

Pedestrians and non-motor vehicles are relatively small in term of size and have a flexible path, and it is difficult to punish illegal acts. Therefore, it is necessary to set pedestrian crossing signals and pedestrian barriers to separate traffic conflicts in the dimension of time and space to regulate the path of pedestrians and cyclists.

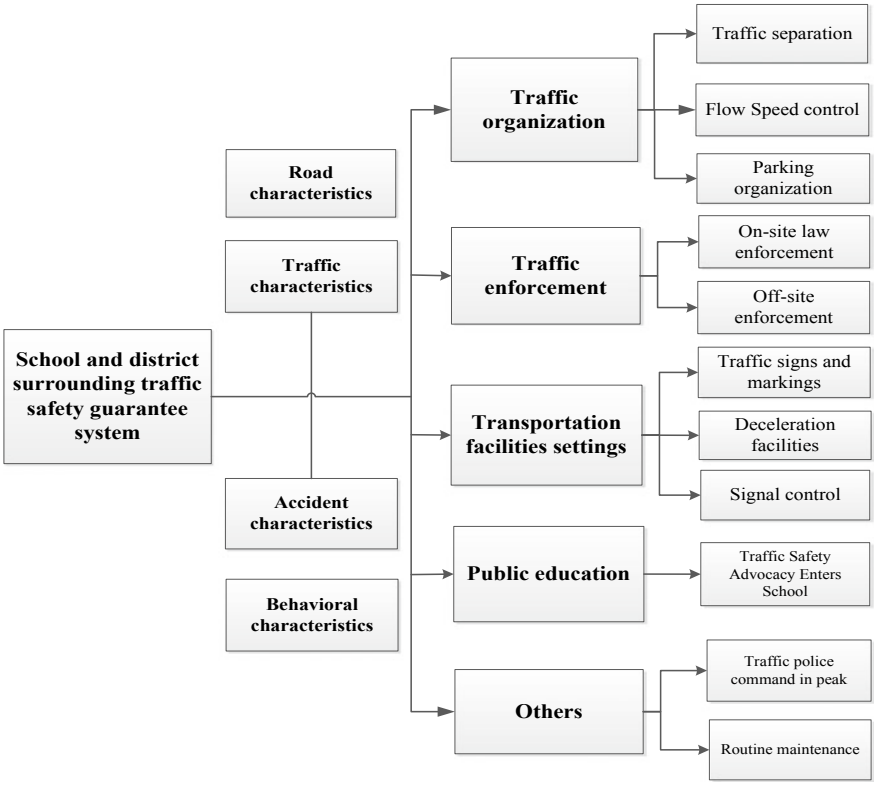


Fig. 1 School and district surrounding traffic safety guarantee system framework

4.2 Transportation Safety Guarantee System

In order to solve the traffic safety problems around the school gate, the paper puts forward the countermeasures of traffic safety management from the four aspects of traffic organization, law enforcement administration, facilities setting and publicity education, and constructs the framework of the traffic safety guarantee system in the campus. As shown in the Fig. 1.

5 Analysis of Transportation Safety Guarantee on the Living Example

Taking Xigong District Experimental Primary School in Luoyang, Henan Province as an example, according to the traffic safety guarantee system, traffic safety management measures for the actual situation of the primary school were formulated.

5.1 Current Investigation of Transportation Safety at School

Luoyang Experimental Elementary School lies in Kaixuan east road, Xigong District, Luoyang city. There is a large number of children who cross the road during the school hours, at the same time there are many vehicles gathering at the entrance of the school to pick up children. According to the traffic characteristics and research needs of the primary school entrance, the study was carried out at the 300 m range of the primary school entrance. As shown in Fig. 2.

(1) Current situation of the road environment. Kaixuan east road for the urban trunk road, two boards for the form, the four lanes, road central segregation, green belts with green belt width is 30 m, inorganic barrier facility, the pavement is 25 m wide, no pedestrian barriers. Only simple front school, pay attention to the pedestrian sign.
(2) Traffic amount is investigated. Making statistics for traffic volume with Kaixuan East road which in front of the experimental school. As shown in the figures of Table 1.

The investigation revealed that the road's vehicle on the two directions is complicated with certain proportions of each type, where the motor vehicle can be reached the largest proportions of 51%, second is bicycle whose share 26%. The smallest

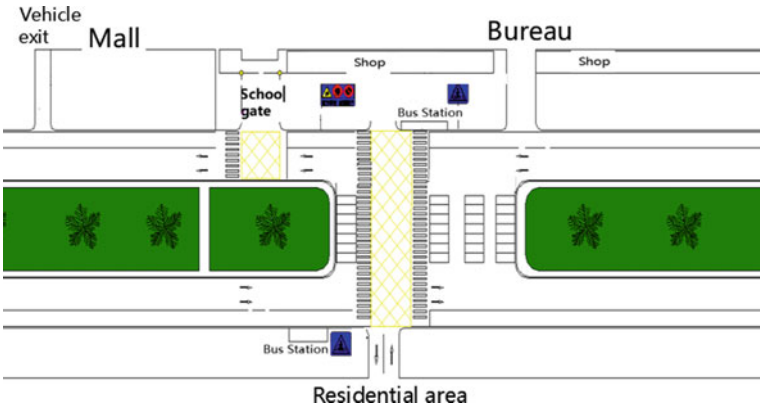


Fig. 2 Current status of the scope of the study

Table 1 Survey of traffic volume

Direction	Number of lanes	Traffic volume (veh/h)		Average speed (km/h)	
		Non-motor vehicle	Motor vehicle	Non-motor vehicle	Motor vehicle
West to east	2	288	696	11.4	14.8
East to west	2	216	516	15.6	30.1

Table 2 Survey of primary school students' cross street behaviors

Type	Counting	Proportion (%)
Compliance with regulations	90	45
Cross the street	76	38
Straight line crossing	34	17

proportion is motorcycle of 6%. Since many parents choose bicycles and mini-cars to pick up children, the proportion of bicycles and mini-cars from east to west is higher than the one from west to east, so total amount is.

(3) The current conditions at the bus stop. The 60 m of the eastern side in the front gate of school is a bus stop, at which there are 7 buses to stop. At the peak, buses are frequent to make sections to inside non-motor cycles to cause congestions and disorder. The full length of the bus stop is 40 m long and can allow two buses to come into the bus stop at the same time. The investigation for pupils' crossing the street is made.
(4) By the investigation, for pupils to cross the street after school, we study the choice pathway for pupils to cross the road and explore the rule for pupils to choose the pathway when they cross the road.

The time of investigation is 15:45–16:15 of March 25, 2017. Taking the type of sampling as an example, which shows the pathway of crossing the street with 200 samples. The result of the investigation is shown in Table 2.

Through the investigation, we found that the population in the front gate of primary school who obeys the laws to cross the street is not beyond 50% while the proportion of crossing the street is large to further cause the disorder at the gate of the school and even cause pedestrian accidents.

5.2 The Analysis of Transportation Safety Problem on Campus

5.2.1 The Aspects of the Transportation Facility

(1) The sidewalk setting is not reasonable. School of the pavement and road median pedestrian access does not connect, in addition, the pavement occupation by non-motor vehicles, majority of people will not choose a pedestrian crossing, but choose to cross the road directly, or illegally cross the straight way. As shown in Fig. 3.

The sidewalk which the western side of experimental school is extremely lowered to use. The non-consistency of facility in crossing street to vehicle and pedestrian were both made the trouble.

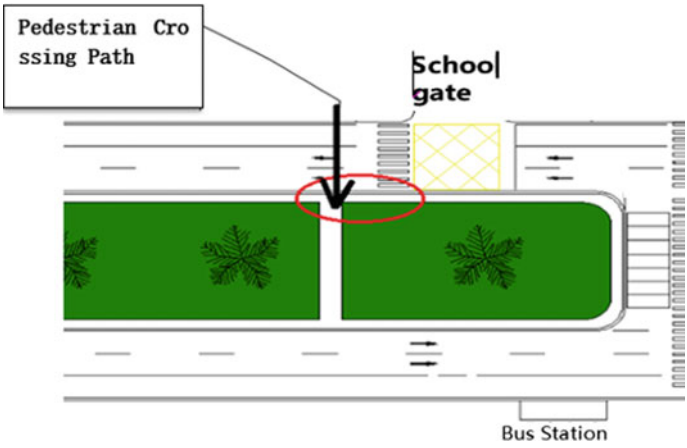


Fig. 3 Crosswalk status

(2) The facility of transportation management is not completed. Through the investigation of current condition around the school:

① The transportation markings are only for pedestrian without non parking markings, limited speed markings or limited speed release markings;
② The facility crossing the street only has sidewalks without preview line sign and pedestrian signals of crossing the street to make pedestrian in a mess when crossing the streets and are difficult to cross the street;
③ No signal and no way to separate conflicted points so that the traffic order at the gate of school cannot get guaranteed. As the key protected area of traffic safety, it is essential to add the signals for guarantee;
④ No fence for pedestrian and no fixed pathway for pedestrian and the phenomena of crossing the street is serious.

5.2.2 The Aspect of Traffic Order

(1) From the result of investigation of vehicle's proportion, the number of cyclists on the road in the west direction is as high as 25%. There is no non-motor vehicle parking area around the school entrance, and non-motor vehicle parking which cause the pedestrian crossing to become "difficult to walk", thus resulting in security risks.
(2) Illegally parking at the road occupies the non-motor driveway. Large amounts of non-motor vehicle flow into motor driveway to bring many hidden trouble to non-motor vehicles and pedestrian.

5.2.3 Bus Stop Setting

(1) Take up non-motorized lanes. The 40 m east of the experimental primary school is a bus station. There are 7 buses in and out of the station. There is no bus shelter, and the traffic flow of the buses is intermixed with the non-motorized traffic flow.
(2) Occlusion of vehicle visibility, rear vehicle parking space is insufficient. After the bus stops, the body directly blocks a pedestrian crossing, and the front of the car is about 1 m in front of the road.

5.3 Transportation Safety Management Solutions

5.3.1 Transportation Organization Methods

(1) Speed-flow control

Set up deceleration facilities. Use the speed limit sign and multiple sets of oscillation mark to remind the driver to slow down and control the speed of the motor vehicle.

(2) Parking organization

For motor vehicles, it is possible to wait in the existing road central green belt parking space, and forbid the roadside parking within 150 m on both sides of the school gate. For non-motorized vehicles, it is necessary to set up the parents' waiting area of 60 m long and 6 m long in front of the mall on the west side of the campus. The parents of non-motor vehicles are waiting in the waiting area.

(3) Add a pedestrian crossing signal. Break up the flashpoint from time to time. It is possible to design the timing scheme of the signal lamp, and to set the phase of the pedestrian crossing.

5.3.2 Traffic Enforcement Methods

(1) On-site law enforcement

Clearing up the illegal parking on the side of the road to guarantee the driving space for non-motor vehicle, at the meantime, strengthening enforcement of vehicle without giving away, and the penalty of causally cross the street of non-motor vehicles and behavior of occupying the zebra crossing.

(2) Off-site law enforcement

The electronic policeman is utilized to monitor the road around the school on the whole day. Avoid parking illegally and running a red light.

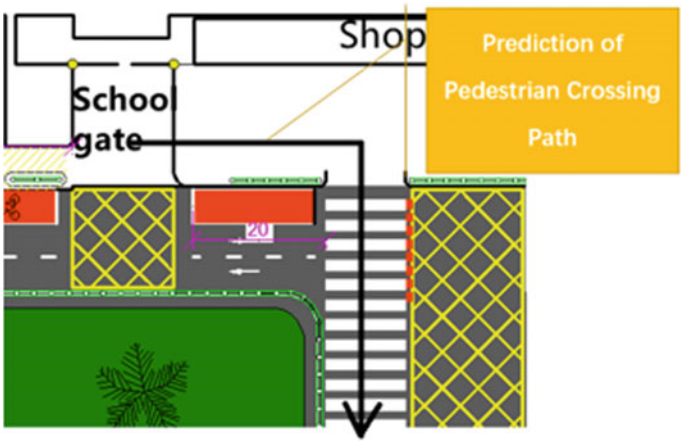


Fig. 4 Schematic of the crosswalk

5.3.3 The Transportation Facility Setting

(1) Set the stop sign, speed limit sign, decelerate slow sign, school sign ahead, and add non-motor lane color road surface. Reorganize the crosswalk to ensure its continuity. To improve the existing pedestrian crossing and add the necessary pedestrian guardrail, the pedestrian crossing path is mandatory, so that the teachers and students can cross the street along the pedestrian crossing to improve the safety. Schematic of crosswalk is shown in the Fig. 4.
(2) Set non-motor vehicle color lanes and reflective stitching to enhance visual identity and ensure the safety of pedestrians crossing the street at night.
(3) The bus station should be designed with the surrounding pedestrian and non-motor vehicle system. The bus station on the north side of the kaide east road should be properly removed from the zebra crossing, and the bus station on the south side of the road should be set to the lower part of the sidewalk to ensure the safety visibility.

5.3.4 Pay Emphasis on Public Education of Transportation Safety

For primary and secondary school students, it is effective and necessary to promote education for traffic safety. Schools can organize various forms of public education, such as invite the police to preach, organize the traffic safety warning film and so on. The traffic safety publicity education was incorporated into the important content of the traffic police department and education department, and effective prevention and control from the source.

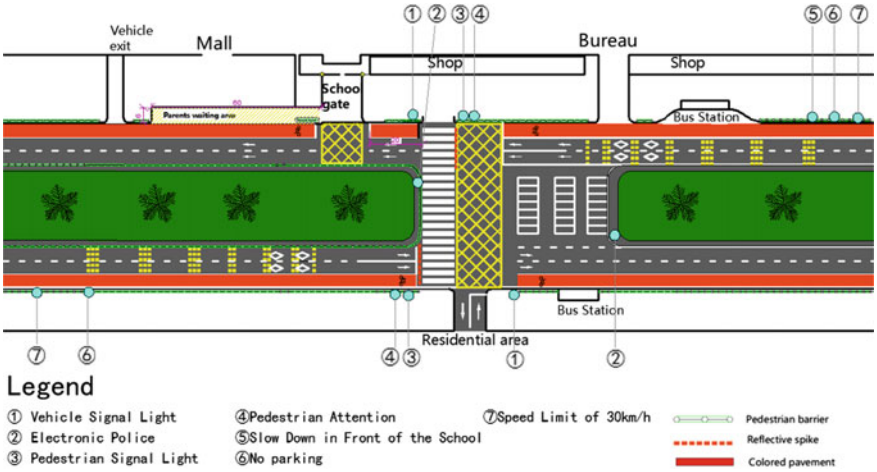


Fig. 5 Implementation of safety and security measures

5.4 Prospect for Implementation of Countermeasures

Facilitating solutions to cross the streets with pedestrian around the experimental primary school are shown in Fig. 5.

Taking the solutions with above safety guarantee:

(1) Utilize the coordination of pavement marking and limited signs to control the speed of the road to make sure safe speed. At the same time, utilizing the combination of Rhombus line and pedestrian signs to make sure the drivers pay attention to the pedestrian;
(2) The pedestrian guardrail is set up on the side of the road, only opening at both ends of the pedestrian crossing, and regulating pedestrian crossing paths;
(3) Adjust the location of the bus station and provide sufficient parking space for rear vehicles;
(4) Set up pedestrian crossing signals to ensure the right of pedestrians crossing the street;
(5) The color lane is added, which visually enables the driver to clear the driving path and reduce the non-conflict of the machine;
(6) Strengthen the law enforcement of education and school gate of primary school students, and reduce the violation of laws and regulations from the subjective and objective aspects. It guarantees the stability of the traffic order around campus and improves the safety of pedestrian crossing.

6 Conclusions

Based on the current situation of Luoyang experimental primary school of the traffic safety field investigation and analysis of traffic data, this paper points out the existing problems of traffic organization, bus stations, traffic order, traffic law enforcement and so forth. In view of the existing traffic safety problems, the security measures of the system are put forward, and the improved traffic environment is evaluated qualitatively. But different schools have their different traffic characteristics, so analysis of specific problems should be taken place case by case instead of being all the same.

References

1. Nakitto MT (2008) Pedestrian traffic injuries among school children in Kawempe, Uganda. Afr Health Sci 8(3):156–159
2. Sun YY, Cong YF, Chai G (2010) Characteristic analysis and countermeasures on elementary and secondary school zone traffic in the county. Traffic Transp
3. Zhao XH, Li JH, Li JF, Rong J (2014) Experimental research on optimal design of school zone traffic safety facilities. J Transp Syst Eng Inf Technol 14(6):207–212
4. Chang XN (2015) Study on optimization of traffic management of urban primary school gate. Master's degree thesis, Jilin University

A Novel Pedestrian Orientation Estimation Method for Autonomous Driving

Ming Gao, LiSheng Jin, Yuying Jiang and Baicang Guo

Abstract Pedestrian orientation estimation is a vital component of autonomous driving system. The challenging factors for pedestrian orientation estimation include pose variations, fast motions, background clutters and crowded people flow. In this paper, we explore a novel pedestrian orientation estimation unified framework, which is based on a monocular camera. Firstly, pedestrian images are normalized to the same size, then extract histogram oriented gradient feature (HOG) which is one of the most effective image descriptor. In addition, we utilize structured Support Vector Machines (SVM) to generate binary classification result. Moreover, Error Correcting Output Coding (ECOC) framework combines with structured SVM to deal with multi-class classification problem. Finally, we conduct our approach on public pedestrian datasets and achieve competitive performance.

Keywords Pedestrian orientation estimation · HOG feature · Structured SVM · ECOC framework

1 Introduction

Pedestrian orientation estimation is a fundamental task for computer vision. It has many applications in related domains, including intelligent transportation system, security surveillance system and autonomous driving system. Since pedestrian orientation estimation has been an important research subject of intelligent transportation system (ITS) for decades [1]. Vision-based advanced driver assistance system (ADAS) uses an on-board monocular camera or stero vision to detect pedestrian, locate pedestrian and estimate pedestrian movement orientation [2, 3]. Furthermore,

M. Gao · L. Jin · B. Guo
College of Transportation, Jilin University, Changchun, China

Y. Jiang (✉)
Department of Ophthalmology, China-Japan Union Hospital of Jilin University, Changchun, China
e-mail: jiangyy@jlu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_34

autonomous driving system (ADS) provides pedestrian orientation estimation to predict and verify pedestrian trajectory. For example, when the pedestrians walk toward the middle of road at unsignalized intersection, autonomous driving system detect pedestrian walk direction then slow down to avoid collision. In this paper, we mainly focus on pedestrian orientation estimation for autonomous driving system.

Pedestrian orientation estimation is a comprehensive problem which combines elements from the computer vision tasks of object detection, where the goal is to identify object classes and localize each using a bounding box, and object tracking, where the goal is to follow the object bounding box with time, and pose estimation, where the goal is to understand relative position of each part and classify with the existing posture. These features allow us to solve this problem from different perspectives. However, pedestrian orientation estimation is challenging because it requires the effective solution for all problem while pedestrian clothing change, pedestrian pose vary, occlusions, and dynamic changes in the background.

The problem of pedestrian orientation estimation has been deeply studied and a large numbers of methods have been published. Existing methods can be roughly divided into two categories: 2D body modeling and 3D body modeling. Early methods are generally 2D model. The 2D body modeling aim to establish a unified model framework to represent entire human body. The methods consider local and holistic image feature have been proposed by Gandhi [4], Shimizu [5] and Goto [6]. Gandhi [4] employed a linear SVM in combination with HMM (Hidden Markov Model) to predict orientation probability. Shimizu [5] proposed a multiple still images to estimate directions. Applying a combination of Haar wavelets and 16 individual SVM classifiers, Goto [6] introduced a Feature Interaction Descriptor (FIND) to use cascade approach to train multi-classifiers. Then, 3D pedestrian orientation methods are inspired by Deformable Part Models (DPM). DPM [7] define a coarse root filter that approximately covers an entire object and higher resolution part filters that cover smaller parts of the object. The model is very suitable for pedestrian detection task. 3D body modeling capture the orientation of body parts, such as the legs, the torso, the arms and the head [8]. One specific paper has been proposed to obtain the face orientation estimation when details such as eyes, nose or even the eyeglasses are detectable [9]. Different characteristic of two methods result in different applicable scenes. The main advantage of 2D model is less computational complexity, which is suitable for real-time traffic object prediction. The 3D model is just the opposite because it gets richer information to achieve higher accuracy, which is suitable for still security monitoring.

In recent years, some novel human orientation detectors have been proposed. These methods consider this problem as a discrete classification problem or a continuous regression problem. Baltieri [10] defined a multi-level HOG feature sets and an array of extremely randomized trees to integrate a mixture of wrapped Gaussian distributions. The result achieve state-of-art performance on benchmarks. Enzweiler [11] proposed an integrated framework for pedestrian classification and pedestrian orientation estimation. The imposed view-specific pedestrian classifiers can train positive and negative samples for orientation estimation task. Similar to Baltieri [10], their method produces continuous prediction by applying the distribution over

the direction as a mixture of Gaussian distribution. Then, joint pedestrian classification and direction estimation is designed by Tao [12]. The work also proposed a new benchmark dataset and a combination of discrete cosine transform with the histogram of orientation gradients.

In this paper, we present a novel pedestrian orientation estimation method. First, our method extract normalized Histogram of Oriented Gradient (HOG) features. HOG feature [13] provide excellent performance at local object appearance and shape, which is one of the most significant descriptors in the field of human detection and some other computer vision task [14–16]. Then, we feed HOG feature to structured Support Vector Machine (S-SVM). Structured SVM [17] is a widely used classifier for object classification in recent years. We design and conduct experiment at baseline pedestrian orientation database. The experiment result shows that our method achieves good performance.

The outline of this paper is as follows. Section 2 defines the classical pedestrian orientation problem. Section 3 explain our HOG feature descriptor combine with S-SVM classifier method. In Sect. 4, we verify proposed method in this paper. Finally, in Sect. 5, we conclude our method and describes some future research issues.

2 Problem Statement

2.1 Problem Statement

In this section, we introduce the pedestrian orientation estimation problem. This problem can be seen as a regression problem or a classification problem. Kota [18] proposed that body orientation was continuous, since artificial discretization of orientation could result in a local optimal performance. Therefore, they used TUD Multiview Pedestrians Dataset but collected continuous annotations manually. In order to specify the body orientation, they move a line segment in a corresponding circle. The user interface is shown in Fig. 1. Although these study would seem to demonstrate that regression methods are more suitable than classification methods, as far as we know this is not the academic mainstream opinion. Despite this regression modeling, there are lots of researchers interested in classification problem. In general, regression problem involve in more time complexity than classification problem. For the purpose of low time cost, we decide to model this task as a classification problem.

We define a series of training data by $\{\mathrm{x}_i, \mathrm{t}_i\}_{i=1}^{N}$, where $x \in R^p$ is an input vector and $t \in R^q$ is a target vector respectively. The main goal of classification problem is to find a function $F(x)$ that can make the expected value of loss function $\Phi(t, F(x))$ minimized as follows:

$$F(x) = \arg\min_{F(x)} \Phi(t, F(x)) \tag{1}$$

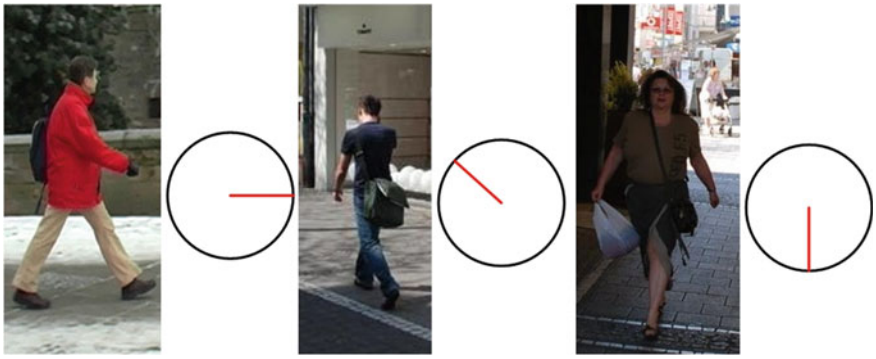


Fig. 1 User interface for continuous orientation annotation

In an attempt to approximate the above expected loss by adding an empirical loss, we use the squared loss function. Equation 1 is reformulated as follows to minimize the sum of squared errors (SSE):

$$F(x) = \arg\min_{F(x)} \sum_{i=1}^{N} \|t_i - F(x_i)\|_2^2 \tag{2}$$

3 Algorithm Design

3.1 HOG Feature

We will focus on low level image features that is basic for our task. In low-resolution face detection tasks, there was an effective image feature called Haar feature which was first proposed by Viola [19] to robustly detect a face. This method uses a boosting cascade of simple features which are based on comparison of region-level intensities. Local Binary Patterns (LBP) feature has also been used as a popular low level image feature. The LBP [20] feature can describe very complicated texture information on the basic of comparisons pixel-level intensity. This hand-crafted feature achieve better perform under the image retrieval task. However, our previous experiments show that these two features are not suitable for our task.

In order to capture richer image feature, we selected a dense feature called Histograms of Oriented Gradient (HOG) to estimate pedestrian orientation. The basic idea is that local object appearance and shape can often be characterized rather well by the distribution of local intensity gradient. The HOG feature allow us to describe the gradients of pixel intensity and contour of objects (see Fig. 2).

Fig. 2 Visualization of HOG feature

Now we explain the basic academic idea of HOG feature. Assuming that the coordinate of one point in an image coordinate is (m, n). Thus the gradient formula is:

$$\begin{aligned} G_m(m,n) &= H(m+1,n) - H(m-1,n) \\ G_n(m,n) &= H(m,n+1) - H(m,n-1) \end{aligned} \tag{3}$$

where $G_m(m,n)$, $G_n(m,n)$ and $H(m,n)$ denote input image pixel (m,n) horizontal gradient, vertical gradient and pixel value. Then, corresponding gradient direction and gradient amplitude can be calculated:

$$\begin{aligned} G(m,n) &= \sqrt{G_m(m,n)^2 + G_n(m,n)^2} \\ \alpha(m,n) &= \tan^{-1}\left(\frac{G_n(m,n)}{G_m(m,n)}\right) \end{aligned} \tag{4}$$

The next step is spatial binning, normalization and descriptor blocks. Notable, the block can be divided as two classes R-HOG and C-HOG stand for rectangular and circular respectively. R-HOG blocks have many similarities to SIFT descriptors [13]. C-HOG focus on a stack of gradient-weight orientation cells [13]. In this paper, we use R-HOG for better experimental results.

3.2 Structured-SVM Classifier

Despite feature extraction plays an important role in a traditional computer vision task, the classifier design module also should pay attention. There are hundreds of classifiers have been published and widely used, such as discriminant analysis, Bayesian, neural networks, support vector machines, decision trees, random forests

[21]. We want to choose a classifier to solve the trade-off between complexity and accuracy. As structured support vector machines have achieved competitive performance, we can highlight four advantages. First, due to the lack of massive pedestrian orientation training samples, support vector machines works better on the small dataset than powerful deep learning methods. Furthermore, support vector machines has solid theoretical basis of statistical learning, perfect form of mathematics and great generalization ability. Then, structured support vector machines is a kind of classification algorithm which can deal with complex outputs like trees, sequences, or sets rather than class labels [22]. The last, structured support vector machines has better discrimination ability than conventional support vector machines. Support Vector Machine was first proposed by Cortes and Vapnik, which also first called linear-SVM [17]. Then, it was improved and developed by plenty of researchers. Structured SVM is one of the most effective structured learning methods.

In order to explain structured SVM clearly, we introduce conventional linear SVM. Assuming that one given training data set $T = \{(\alpha_1, \beta_1), (\alpha_2, \beta_2), (\alpha_3, \beta_3), \ldots, (\alpha_N, \beta_N)\}$, where $\alpha_1 \in X = R^n$, $\beta_i \in Y = \{-1, +1\}$, $i = 1, 2, \ldots, N$, α_i is the ith feature vector, β_i is α_i class label respectively. When $\beta_i = +1, \alpha_i$ is positive example; when $\beta_i = -1, \alpha_i$ is negative example. The main idea of linear SVM is to map the input vectors into some high dimensional feature space Z. In order to solve linearly separable train dataset by finding maximum-margin hyperplane (see Fig. 3):

$$w^* \cdot x + b = 0 \tag{5}$$

where w^* is the weight, x represent support vectors, b is the intercept. And the corresponding linear decision function $I(z)$ in the feature space can be formed as:

$$f(x) = sign(w^* \cdot x + b) \tag{6}$$

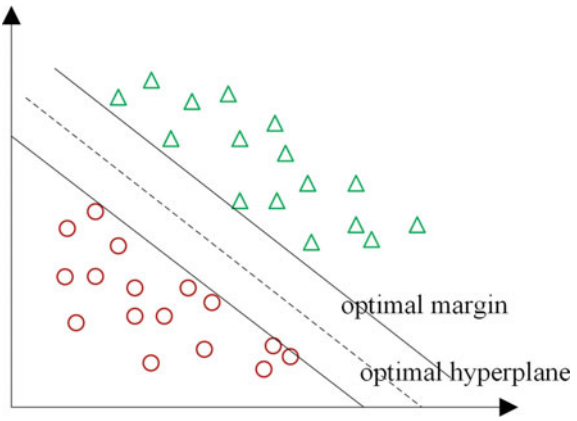


Fig. 3 Visualization of linear SVM

Different from common classifiers, structured SVM learn a prediction function $f:\chi \rightarrow \gamma$ to directly estimate the direction sequences. The output space can be seen as the space of all transformations γ instead of common binary labels $+1$ and -1. In this paper, a pair (x, y) is a labeled example where y is the desired direction of the target. The structured SVM framework which introduces a new discriminant function $F:\chi \times \gamma \rightarrow \mathbb{R}$ that can be used for prediction according to:

$$y_t = f(\mathrm{x}_t) = \arg\min_{\mathrm{y}\in\gamma} F(\mathrm{x}_t, \mathrm{y}) \tag{7}$$

F measures the relationship between (x, y) pairs, and gives a higher score to those which are more similar. Then we add restrict to this expression $F(\mathrm{x}, \mathrm{y}) = \langle w, \phi(\mathrm{x}, \mathrm{y})\rangle$, where $\phi(\mathrm{x}, \mathrm{y})$ denote a joint kernel map, it can be learned from a given set of example pairs $\{(\mathrm{x}_1, \mathrm{y}_1), (\mathrm{x}_2, \mathrm{y}_2), \ldots, (\mathrm{x}_n, \mathrm{y}_n)\}$ by solving the convex objective function:

$$\begin{aligned} \min_{\mathrm{w}} \quad & \frac{1}{2}\|\mathrm{w}\|^2 + \mathrm{C}\sum_{i=1}^{n}\xi_i \\ \text{s.t.} \quad & \forall i{:}\xi_i \le 0 \\ & \forall i, \forall \mathrm{y} \ne \mathrm{y}_i : \langle \mathrm{w}, \delta\phi_i(\mathrm{y})\rangle \ge \Delta(\mathrm{y}_i, \mathrm{y}) - \xi_i \end{aligned} \tag{8}$$

where $\delta\phi_i(\mathrm{y}) = \phi(\mathrm{x}_i, \mathrm{y}) - \phi(\mathrm{x}_i, \mathrm{y})$. This loss function plays an important role in this paper which can make sure that all samples being treated equally.

To solve (8), we adopt the approach proposed by Bordes [23, 24]. A standard Lagrangian duality method can be used to transform its corresponding dual form:

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i,\mathrm{y}\ne\mathrm{y}_i} \Delta(\mathrm{y}, \mathrm{y}_i)\alpha_i^{\mathrm{y}} - \frac{1}{2} \sum_{\substack{i,\ \mathrm{y}\ne\mathrm{y}_i \\ j,\ \overline{\mathrm{y}}\ne\mathrm{y}_j}} \alpha_i^{\mathrm{y}}\alpha_j^{\overline{\mathrm{y}}}\langle\delta\phi_i(\mathrm{y}), \delta\phi_i(\overline{\mathrm{y}})\rangle \\ \text{s.t.} \quad & \forall i, \forall \mathrm{y} \ne \mathrm{y}_i{:}\alpha_i^{\mathrm{y}} \ge 0 \\ & \forall i : \sum_{\mathrm{y}\ne\mathrm{y}_i} \alpha_i^{\mathrm{y}} \le C \end{aligned} \tag{9}$$

The discriminant function also could be rewrote as $F(\mathrm{x}, \mathrm{y}) = \sum_{i,\overline{\mathrm{y}}\ne\mathrm{y}_i} \alpha_i^{\overline{\mathrm{y}}}\langle\delta\phi_i(\overline{\mathrm{y}}), \phi(\mathrm{x}, \mathrm{y})\rangle$. The joint kernel map $\phi(\mathrm{x}, \mathrm{y})$ only occurs inside inner products, it can be defined an appropriate joint kernel function $k(\mathrm{x}, \mathrm{y}, \overline{\mathrm{x}}, \overline{\mathrm{y}}) = \langle\phi(\mathrm{x}, \mathrm{y}), \phi(\overline{\mathrm{x}}, \overline{\mathrm{y}})\rangle$. The kernel functions we use during estimating orientation is radial basis function.

As in [21], by adjusting the parameters in (9) according to

$$\beta_i^{\mathrm{y}} = \begin{cases} -\alpha_i^{\mathrm{y}} & \mathrm{y} \ne \mathrm{y}_i \\ \sum\limits_{\overline{\mathrm{y}}\ne\mathrm{y}_i} \alpha_i^{\overline{\mathrm{y}}} & \text{otherwise} \end{cases} \tag{10}$$

the dual form can be simplified to

$$\max_{\beta} \quad -\sum_{i,\mathrm{y}} \Delta(\mathrm{y}, \mathrm{y}_i)\beta_i^{\mathrm{y}} - \frac{1}{2}\beta_i^{\mathrm{y}}\beta_j^{\overline{\mathrm{y}}}\left\langle \phi(\mathrm{x}_i, \mathrm{y}), \phi(\mathrm{x}_j, \overline{\mathrm{y}})\right\rangle$$
$$\text{s.t.} \quad \begin{array}{l} \forall i, \forall \mathrm{y}: \beta_i^{\mathrm{y}} \geq \delta(\mathrm{y}, \mathrm{y}_{\mathrm{i}})C \\ \forall i : \sum_{\mathrm{y}} \beta_i^{\mathrm{y}} = 0 \end{array} \tag{11}$$

where $\delta(\mathrm{y}, \overline{\mathrm{y}}) = 1$ if $\mathrm{y} = \overline{\mathrm{y}}$ and $\delta(\mathrm{y}, \overline{\mathrm{y}}) = 0$ if $\mathrm{y} \neq \overline{\mathrm{y}}$.

The SMO (Sequential Minimal Optimization) [25] step play a core role which improves (11) corresponding pairs of coefficients $\beta_i^{\mathrm{y}_+}$ and $\beta_i^{\overline{\mathrm{y}}_-}$. The next step of the online learning algorithm focus on how to choose the triplet $(i, \mathrm{y}_+, \mathrm{y}_-)$ which could be optimized by SMO step. Assuming that we have a given triplet $(i, \mathrm{y}_+, \mathrm{y}_-)$ to solve the feasible search direction with the highest gradient, where gradient of (11) with respect to one single coefficient β_i^Y which is given by

$$\begin{aligned} g_i(\mathrm{y}) &= -\Delta(\mathrm{y}, \mathrm{y}_i) - \sum_{j,\overline{\mathrm{y}}} \beta_j^{\overline{\mathrm{y}}}\left\langle \phi(\mathrm{x}_i, \mathrm{y}), \phi(\mathrm{x}_j, \overline{\mathrm{y}})\right\rangle \\ &= -\Delta(\mathrm{y}, \mathrm{y}_i) - F(\mathrm{x}_i, \mathrm{y}) \end{aligned} \tag{12}$$

3.3 Error Correcting Output Coding

Above all, we can apply structured support vector machines to solve binary class problem. In fact, we need to convert a binary classifier to multiclass classifiers for the purpose of structured support vector machines which can not deal with multi-class problem. There are three main strategies for this problem: One vs One (short for OvO), One vs Rest (short for OvR), Many vs Many (short for MvM). In this paper, we utilize the classical MvM strategy. We adopt Error Correcting Output Coding (ECOC) method [26–28]. The ECOC method can be divided by two parts: coding and decoding [28]. In the coding phase, we use L categories to divide to M classes, each partition can result some positive examples and some negative examples, which can combine a two classes training dataset. Then, we have a total of Mtraining set and classifiers. In the decoding phase, considering the Hamming distance between two words of code:

$$D_H(c_i, c_j) = \sum_{h=1}^{L} \left| c_{ih} - c_{jh} \right| \tag{13}$$

A decoding phase is applied at the output of the separate and one sample is distributed to the respect class if the Hamming distance between the output word

o(x) and the k-th codeword is minimum:

$$\omega_k = \arg\min_h D_H(c_h, o) \tag{14}$$

The above ECOC method can be summarized as a coding matrix. The row is a L-length binary string and the column is Mdifferent classes as shown in Table 1.

Table 1 Eight classes coding matrix

	W1	W2	W3	W4	W5	W6	W7	W8
F1	1	1	0	0	0	0	0	0
F2	1	0	1	0	0	0	0	0
F3	1	0	0	1	0	0	0	0
F4	1	0	0	0	1	0	0	0
F5	1	0	0	0	0	1	0	0
F6	1	0	0	0	0	0	1	0
F7	1	0	0	0	0	0	0	1
F8	0	1	1	0	0	0	0	0
F9	0	1	0	1	0	0	0	0
F10	0	1	0	0	1	0	0	0
F11	0	1	0	0	0	1	0	0
F12	0	1	0	0	0	0	1	0
F13	0	1	0	0	0	0	0	1
F14	0	0	1	1	0	0	0	0
F15	0	0	1	0	1	0	0	0
F16	0	0	1	0	0	1	0	0
F17	0	0	1	0	0	0	1	0
F18	0	0	1	0	0	0	0	1
F19	0	0	0	1	1	0	0	0
F20	0	0	0	1	0	1	0	0
F21	0	0	0	1	0	0	1	0
F22	0	0	0	1	0	0	0	1
F23	0	0	0	0	1	1	0	0
F24	0	0	0	0	1	0	1	0
F25	0	0	0	0	1	0	0	1
F26	0	0	0	0	0	1	1	0
F27	0	0	0	0	0	1	0	1
F28	0	0	0	0	0	0	1	1

3.4 Class Imbalance Problem

In regression and classification training step, typically every training sample is treated equally. However, in real world, we collect training data in a specific scene, for example, we collect pedestrian images at intersection, and pedestrian orientation is more likely parallel to the direction of the crosswalk. The training data are not uniform distribution. Data imbalances can lead to worse classification result which data are more likely to be classified into majority classes. In order to solve data imbalance problem, we employed the SMOTE method proposed by Chawla [29]. The algorithm is based on the assumption that near distance minority samples are still the same class, which are used to control the generation and distribution of artificial samples to achieve data equalization. SMOTE algorithm can be described in Algorithm 1:

Algorithm SMOTE(T,N,k)

Input: Number of minority class samples T ;Amount of SMOTE N %;Number of nearest neighbors k

Output: $(N/100)*T$ synthetic minority class samples

1. (If N is less than 100%, randomize the minority class samples as only a random percent of them will be SMOTE.)
2. if $N<100$
3. then Randomize the T minority class samples
4. $T=(N/100)*T$
5. $N=100$
6. endif
7. N=(int)$(N/100)$ (The amount of SMOTE is assumed to be in integral multiples of 100.)
8. $k=$ Number of nearest neighbors
9. *numattrs* =Number of attributes
10. *Sample*[][] : array for original minority class samples
11. *newindex* : keeps a count of number of synthetic samples generated, initialized to 0
12. *Synthetic*[][] : array for synthetic samples

(Compute k nearest neighbors for each minority class sample only.)

13. for $i \leftarrow 1$ to T
14. Compute k nearest neighbors for i , and save the indices in the *nnarray*
15. Populate $(N, i, nnarray)$
16. endfor

Populate$(N, i, nnarray)$ (Function to generate the synthetic samples)

17. while $N \neq 0$
18. Choose a random number between 1 and k, call it nn. This step chooses one of the k nearest neighbors of i.
19. for $attr \leftarrow 1$ to $numattrs$
20. Compute: $dif = Sample[nnarray[nn]][attr] - Sample[i][attr]$
21. Compute: gap = random number between 0 and 1
22. $Synthetic[newinder][attr] = Sample[i][attr] + gap * dif$
23. endfor
24. $newindex++$
25. $N = N - 1$
26. endwhile
27. return(End of Populate)

The method can effectively solve the class imbalance problem.

3.5 Dimension Reduction

All machine learning process may involve in high dimensional computational complexity problem, which also called 'curse of dimensionality'. In order to solve this problem, plenty of dimensionality reduction methods have been published. Principal component analysis (PCA) is one of the most popular and effective approaches. By this method, the mapping of high-dimensional space in low-dimensional subspace can be obtained. PCA [30] algorithm can be described in Algorithm 2:

Algorithm PCA
Input: Sample dataset $D = (x_1, x_2, \ldots, x_m)$;
The dimension of low–dimensional space d
Output: Projection matrix $W = (w_1, w_2, \ldots, w_d)$

1. Centralized all samples: $x_i \leftarrow x_i - \frac{1}{m}\sum_{i=1}^{n} x_i$
2. Calculate the sample covariance matrix XX^T
3. Decompose eigenvalue of the sample covariance matrix XX^T
4. Get the largest d eigenvalues corresponding to the eigenvectors
5. end

4 Experiments and Analysis

The experiments are implemented by MATLAB with Classification Learner. All experiments are conducted on a PC with Intel Core i7-7700 CPU (3.60 GHz/16 GB RAM) and a single NVIDIA GeForce 1080Ti GPU.

We divide the pedestrian orientation as eight orientations. Figure 4 shows our sample images of eight orientations labeled 1–8. We define every orientation as follows: the label 1 denotes the body exactly opposite the camera, while the label 5 denotes the body faces the camera. Then, each label is 45° away from adjacent labels.

For a given image of pedestrian I, we need to resize the input image to 48×96 pixels for normalized purpose and calculate the HOG feature. After several tests, we use 8×8 pixel cell, 2×2 cell block, and 9 quantized orientations HOG feature. The pixel gradients are voted into 9 orientations bins in 0–180°, described in [13]. Notably, every HOG descriptor have 2304 elements from a single pedestrian image.

We use ECOC-SSVM to train body direction. Because our labels are eight classes. Table 1 shows our 8 classes coding matrix where {W1, W2, …, W8} means that the 8 orientation classes, {F1, F2, …, F28} means that the 28 classification comparison. When implementing above multi-class structured SVM from the training dataset, we utilized LIBSVM [31] and MATLAB Classification Learner to train and predict our model.

We conduct experiments on the PDC [15] benchmark dataset which consists of 11,562 images of pedestrians with orientation annotations (see Fig. 5). Note that the size of the dataset is two times larger than that TUD dataset [12]. The training set contains all the 11,562 images. Then we utilize k-fold cross-validation to validate our model and set k value as 5. All of the images are gray images. And all of the images in this dataset are obtained "in the wild". The dataset contain different poses and clothing, making the dataset much more challenging.

We select two main performance evaluation metrics. Mean absolute error (MAE) of angular distance is selected as the first one:

$$d_{angle}(t_1, t_2) = \min_{k \in \{0,-1,+1\}} |t_1 - t_2 + 360k| \tag{15}$$

Fig. 4 Body orientation representation

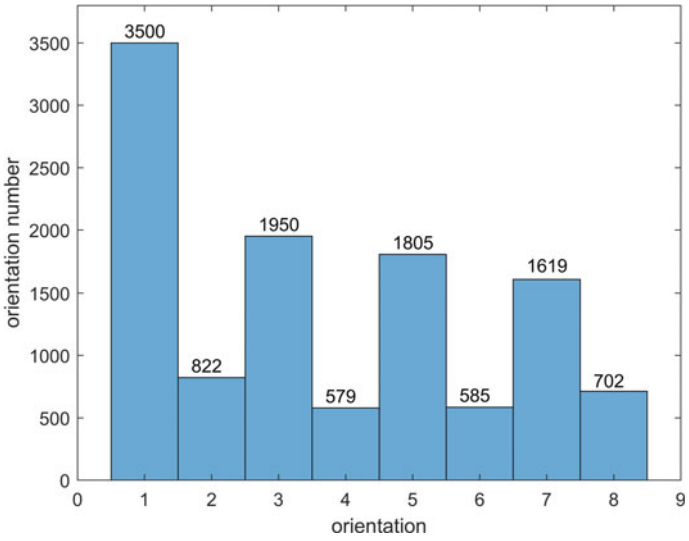


Fig. 5 PDC dataset orientation number

where t_1 means estimation value and t_2 means true value. Then the second measures was Accuracy which defined as the percentage of right prediction from the ground truth. We think that Accuracy is more effective than MAE in practice, because large errors can strongly influence the MAE.

In order to obtain better classification results, we compared different kernel functions. Additionally, we explored the influence of dimensionality reduction method. Table 2 demonstrates the Accuracy and MAE of the proposed method. We adapt different kernel function including linear, quadratic, cubic, fine gaussian, medium gaussian and coarse gaussian. Note that the kernel scale of fine gaussian, medium gaussian and coarse gaussian is 12, 48 and 190 respectively. And we also evaluate the performance based on PCA. The dimension of HOG feature is 2304 and under the circumstance of 95% variance PCA is 826 dimension. In general, without PCA methods are slightly better than the same kernel function with PCA, because of the PCA lose some features inevitably. Quadratic kernel function performs favorably among all kernel function, which achieve 81.6 and 84.2% accuracy. Linear and cubic kernel function achieve better accuracy. The gap between different gaussian kernel function is obvious. As medium gaussian kernel achieve competitive performance, fine gaussian kernel achieve worse performance of 30.7, because a small number of kernel scale limit the performance of classifier. Table 2 also lists a MAE comparison of our methods. There is no doubt that quadratic kernel function without PCA is superior other methods, which MAE is 28.1. It is notable that the performance of coarse gaussian method with or without PCA is much of different. In case of original features, coarse gaussian achieve better performance while coarse gaussian achieve much worse result with the PCA method. Overall, the accuracy is negatively related

Table 2 Result of pedestrian orientation estimation

Method	Kernel function	Accuracy (%)	MAE (°)
HOG + SVM (with PCA)	Linear	78.7	39.2
	Quadratic	81.6	32.9
	Cubic	81.6	33.5
	Fine Gaussian	30.7	129.7
	Medium Gaussian	79.3	36.8
	Coarse Gaussian	31.3	126.8
HOG + SVM (without PCA)	Linear	81.0	35.2
	Quadratic	84.2	28.1
	Cubic	84.0	38.7
	Fine Gaussian	30.4	129.7
	Medium Gaussian	84.1	28.6
	Coarse Gaussian	73.1	48.5
Human [15]	–	90.7	9.1

to MAE. Compared with human, our method has a similar accuracy while human MAE is much smaller than our method.

We further analyze the detail of classification performance under different methods. Figures 6 and 7 show the confusion matrix of different kernel function orientation estimation. Figure 6 is confusion matrix with PCA. Figure 7 is confusion matrix without PCA. The row denotes the ground truth orientation label and the column denotes the predicted orientation label. All the methods perform best in the orientation of pedestrian opposite the camera which encode number 1 that has the accuracy larger than 94%. Because encode number 1 has the most samples. The integer direction perform better results than oblique direction. The oblique encode number 2, 4, 6, 8 are more likely to be classify to adjacent odd number. It is notable that medium gaussian kernel achieve better performance than other gaussian kernel under all the situation. Despite awful coarse gaussian kernel with PCA, coarse gaussian kernel without PCA achieve performance better between polynomial kernel and other gaussian kernel. In a summary, quadratic kernel function is the most effect method under the accuracy index.

Figures 8 and 9 show the ROC graphs. The main advantage of ROC graphs is that they enable visualizing and organizing classifier performance without regard to class distributions or error costs [32]. AUC (Area under an ROC curve) is a single scalar value calculate the area under the real ROC curve, which represent statistical property of the classifier. Higher AUC means better result. EER (equal error rate) is the point on the ROC curve that corresponds to have an equal probability of miss-classifying a positive or negative sample. This point is obtained by intersecting the ROC curve with

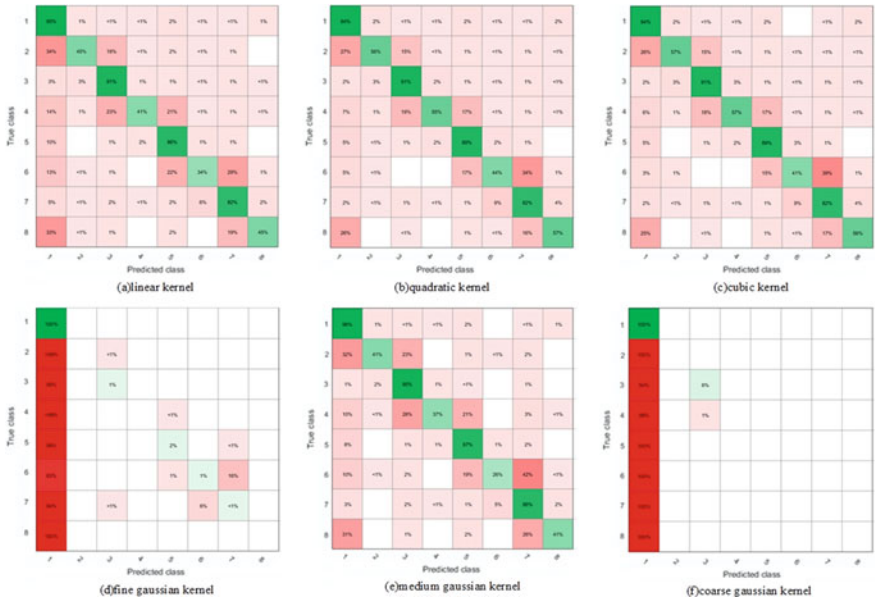


Fig. 6 Confusion matrix with PCA

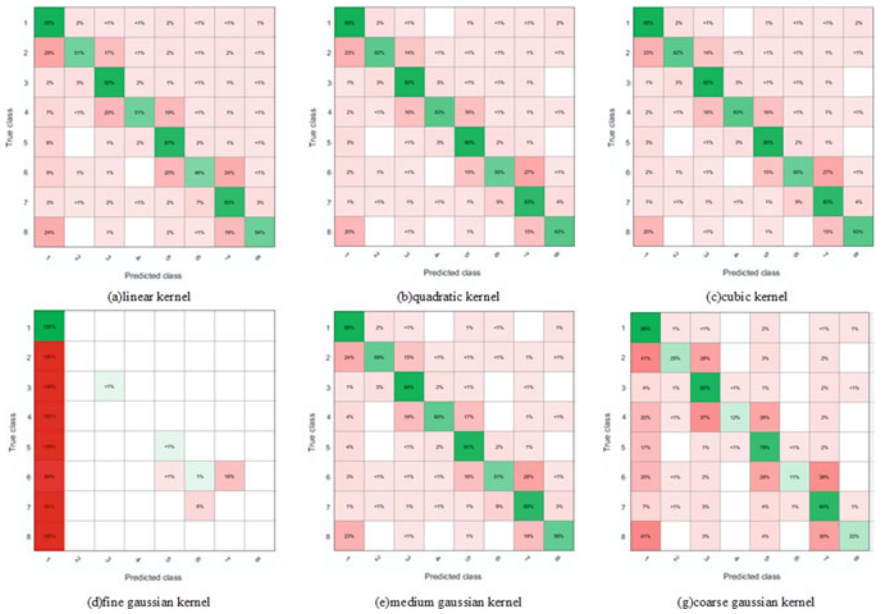


Fig. 7 Confusion matrix without PCA

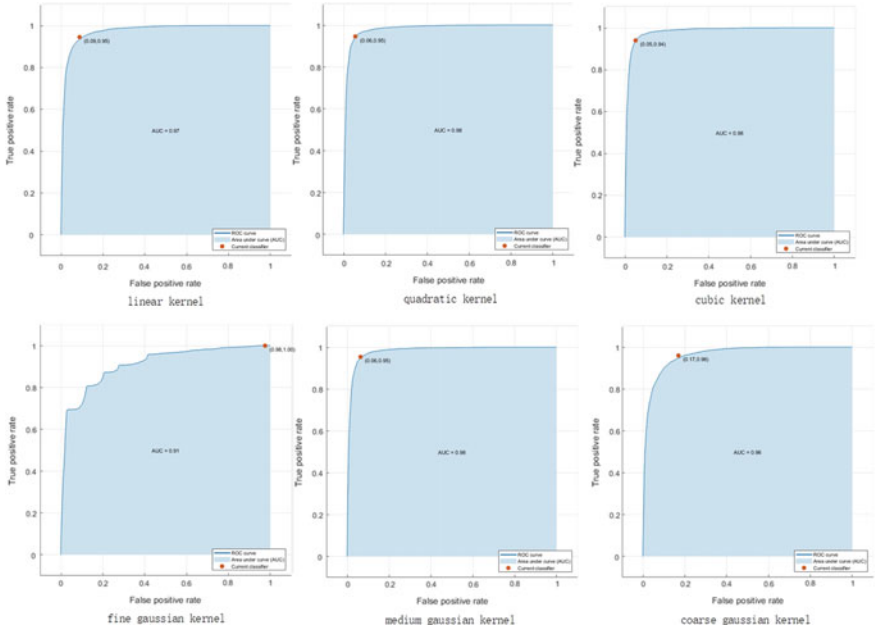


Fig. 8 ROC curve of orientation estimation (with pca)

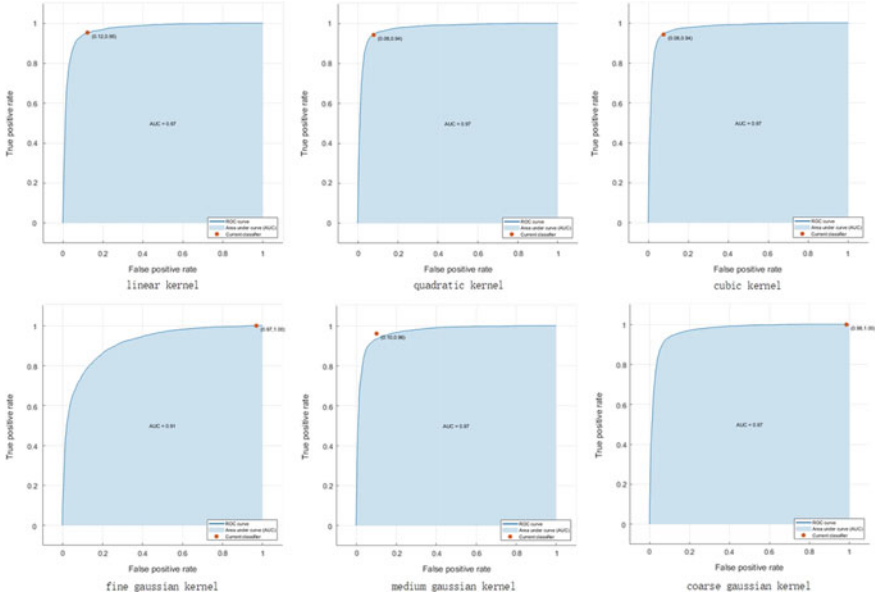


Fig. 9 ROC curve of orientation estimation (without pca)

Table 3 Synthesize pedestrian orientation image number

Label	2	4	6	8
Image number	1644	1158	1170	1404

a diagonal of the unit square. Figure 8 is classifiers with PCA and Fig. 9 is classifiers without PCA. In each subgraph, the shape of ROC is like a round rectangle indicate that the experiment results are relatively good. We can find that all the AUC is larger than 0.95 except fine gaussian kernel with PCA which the AUC is 0.9. Higher AUC value means that the S-SVM classifiers achieve better overall performance. There is no obvious difference between PCA and without PCA methods. It is shown that quadratic kernel without PCA reach the highest AUC value which consistent with confusion matrix. And the EER of above method is (0.06, 0.95), which means that the performance is competitive well.

In order to ease distribution unbalanced problem, we try to use SMOTE algorithm to generate artificial orientation features. We set the SMOTE parameter nearest neighbors to $k = 5$. We synthesize 4 less number direction features as Table 3, the new dataset consists of 14,250 image features. As we have proven that gaussian kernel function was not very suitable for our task, we just test the linear and polynomial kernel.

Table 4 shows the new experiment results. We can find that cubic kernel without PCA achieve the best accuracy and MAE. All the results slighty better than original experiment, which 4% accuracy higher than before. MAE is also decrease about 6%, which suggest that SMOTE is very effective. Above all, we can conclude that structured SVM with polynomial kernel is more suitable for pedestrian orientation estimation.

Table 4 Pedestrian orientation estimation with SMOTE

Method	Kernel function	Accuracy (%)	MAE (°)
HOG + SVM (with PCA)	Linear	82.2	33.0
	Quadratic	86.7	25.1
	Cubic	88.3	21.0
HOG + SVM (without PCA)	Linear	85.1	37.6
	Quadratic	87.9	22.9
	Cubic	89.0	20.6
Human [15]	–	90.7	9.1

5 Conclusions

A novel approach for pedestrian orientation estimation was proposed. We have shown that an ECOC framework structured SVM was suitable for multi-class classification problem. The experimental results denote that the excellent performances in human orientation estimation are achieved by our presented method based on popular PDC database. Different from time-consuming deep learning method, the main advantage of our method is computational simplicity. The experimental results showed the effectiveness of our human orientation estimation model. In the future work, we plan to explore feature fusion method to encode robust classifier and pedestrian occlude should be discussed.

References

1. Zhao G, Takafumi M, Shoji K (2012) Video based estimation of pedestrian walking direction for pedestrian protection system. J Electron 29(1):72–81
2. Geronimo D, Lopez AM (2013) Vision-based pedestrian protection systems for intelligent vehicles. Springer Publishing Company, Incorporated, New York
3. Böhmländer D, Doric I, Appel E et al (2013) Video camera and capacitive sensor data fusion for pedestrian protection systems. In: Intelligent solutions in embedded systems. IEEE, pp 1–7
4. Gandhi T, Trivedi MM (2008) Image based estimation of pedestrian orientation for improving path prediction. In: Intelligent vehicles symposium. IEEE, pp 506–511
5. Shimizu H, Poggio T (2004) Direction estimation of pedestrian from multiple still images. In: Intelligent vehicles symposium. IEEE, pp 596–600
6. Goto K, Kidono K, Kimura Y et al (2011) Pedestrian detection and direction estimation by cascade detector with multi-classifiers utilizing feature interaction descriptor. In: Intelligent vehicles symposium. IEEE, pp 224–229
7. Forsyth D (2010) Object detection with discriminatively trained part-based models. IEEE Trans Pattern Anal Mach Intell 1627–1645
8. Andriluka M, Roth S, Schiele B (2010) Monocular 3D pose estimation and tracking by detection. In: Computer vision and pattern recognition. IEEE, pp 623–630
9. Setthawong P, Vannija V (2012) Improving the estimation of head pose orientation: by using eyeglasses as a key feature. In: international conference on information technology and multimedia. IEEE, pp 1–6
10. Baltieri D, Vezzani R, Cucchiara R (2012) People orientation recognition by mixtures of wrapped distributions on random trees. In: Computer vision—ECCV 2012. Springer, Berlin, Heidelberg, pp 270–283
11. Enzweiler M, Gavrila DM (2010) Integrated pedestrian classification and orientation estimation. In: Computer vision and pattern recognition. IEEE, pp 982–989
12. Tao J, Klette R(2014) Part-based RDF for direction classification of pedestrians, and a benchmark. In: Computer vision—ACCV 2014 workshops. Springer International Publishing, New York, pp 418–432
13. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: IEEE computer society conference on computer vision & pattern recognition. IEEE Computer Society, pp 886–893
14. Bratanič B, Pernuš F, Likar B et al (2014) Real-time rotation estimation using histograms of oriented gradients. PloS One 9(3):e92137
15. Liu CH, Lin JK (2015) Integral histogram with random projection for pedestrian detection. PloS One 10(11):e0142820

16. Zhang M, Ming X (2016) Human detection using random color similarity feature and random ferns classifier. PloS One 11(9):e0162830
17. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
18. Hara K, Chellappa R (2017) Growing regression tree forests by classification for continuous object pose estimation. Kluwer Academic Publishers, The Netherlands
19. Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput Vision 57(2):137–154
20. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
21. Cernadas E, Amorim D (2014) Do we need hundreds of classifiers to solve real world classification problems? J Mach Learn Res 15(1):3133–3181
22. Tsochantaridis I, Joachims T, Hofmann T (2005) Large margin methods for structured and interdependent output variables. J Mach Learn Res 6(2):1453–1484
23. Bordes A, Gallinari P, Weston J (2007) Solving multiclass support vector machines with LaRank. ACM, pp 89–96
24. Bordes A, Usunier N, Bottou L (2008) Sequence labelling SVMs trained in one pass. In: European conference on machine learning and knowledge discovery in databases. Springer, Berlin, Germany, pp 146–161
25. Platt JC (1999) Fast training of support vector machines using sequential minimal optimization, advances in kernel methods. MIT Press, Cambridge, MA, pp 185–208
26. Dietterich TG, Bakiri G (2012) Solving multiclass learning problems via error-correcting output codes. J Artif Intell Res 2(1):263–286
27. Pujol O, Radeva P, Vitrià J (2006) Discriminant ECOC: a heuristic method for application dependent design of error correcting output codes. IEEE Trans Pattern Anal Mach Intell 28(6):1007–1012
28. Simeone P, Tax DMJ, Duin RPW et al (2008) A fast approach to improve classification performance of ECOC classification systems. In: Structural, syntactic, and statistical pattern recognition. Springer, Berlin, Heidelberg, pp 459–468
29. Chawla NV, Bowyer KW, Hall LO (2002) SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 16(1):321–357
30. Tipping ME, Bishop CM (1999) Probabilistic principal component analysis. J Roy Stat Soc 61(3):611–622
31. Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. ACM, p 1
32. Fawcett T (2006) An introduction to ROC analysis. Pattern Recogn Lett 27(8):861–874

Operation Optimization Considering Order Cancellation and Ticket Discount for On-Demand Bus System

Haipeng Shao, Xingying Chen, Yuxuan Wang and Sufeng Wu

Abstract The emergence of multi-source and mass travel data and the gradual maturity of acquisition methods make the realization methods of travel demand more diversified. The on-demand bus is a pattern that can meet the needs of personalized and high-quality travel. This project proposes an on-demand bus system based on response to random users' real-time requests. Under the condition of fixed origin points and destination points, taking the minimum total waiting time of passengers and the maximum profit of the bus as the goals, a model based on the variables of bus fare discount and request valid time is established, and LINGO and the actual car-hailing data are used to calculate and check the model. Through calculation results and the comparison with the traditional bus and the express, the on-demand bus system has shorter travel time than the traditional bus and costs less than the express. Thus the research can be thought to provide a scientific basis for the establishment of the on-bus bus system, and provide theoretical support for further research.

Keywords On-demand bus · Waiting time of passengers · Ticket price compensation mechanism

H. Shao (✉) · Y. Wang
Chang'an University, Middle Section of Nan Erhuan Road, Xi'an City 710064, Shannxi Province, China
e-mail: shaohp@chd.edu.cn

Y. Wang
e-mail: 869748533@qq.com

X. Chen · S. Wu
Institute of Transportation Development Strategy & Planning of Sichuan Province, Chengdu City 610041, Sichuan Province, China
e-mail: chenxingying@chd.edu.cn

S. Wu
e-mail: wusufeng@126.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_35

1 Introduction

With the rapid development of the economy, the level of motorization in Chinese cities has been greatly promoted. To satisfy the increasing travel demand of residents and alleviate the problem of urban traffic congestion, after the establishment of the strategic goal of bus priority development, the local government has promulgated a series of policies and measures, such as expanding the layout of the public transportation network and strengthening the construction of rail transit. These measures ease the traffic pressure in the city to a certain extent, but still can't completely meet the residents' diversified travel demands, especially in the increasingly high requirements on life quality today. People focus more on the travel experience of minority, individualization and high-quality. Meanwhile the accuracy, accessibility, diversity and comfort of travel gradually get more attention.

With the extensive use of smart phone App based on the LBS (location based service), it is possible to obtain the user's mobile terminal location information, which provides technical support for the rapid development of express cars, premier cars and car sharing. But the high travel costs of express cars and premier cars are unaffordable for most people. Therefore, it is imperative to develop an on-demand bus, whose travel experience is more comfortable than the traditional bus, and the travel cost is lower than the express cars. The on-demand bus is a new travel mode based on the internet background. It provides a new idea for improving the quality of the bus transit system, and it is a supplement to the low quality of the existing public transport service.

Liu [1] compared the passenger attribute, line stations, bus fare, service quality and other aspects of on-demand bus and the traditional bus, established site and route planning model and ticket pricing model, laying a theoretical foundation for further on-demand bus research. Guo [2] analyzed the key elements of the on-demand bus system and established the on-demand bus model. His research provided the basis to the on-demand bus system in the sparsely populated area. Li [3] built Logistic model of on-demand bus fare, and he studied the ticket ranges that commuters could accept, the study pointed out that whether there were seats, the number of bus stops along the way and commuting distance were important factors affecting the city residents receive an on-demand bus. Huang et al. [4] proposed a new nonlinear bus fare structure based on distance, which was measured by Euclidean distance between the destination and the destination. The results indicated the advantage over existing fare structures: it reflected the 'true cost' of a passenger's trip. Jinux [5] proposed a methodology for the optimal design of a suburban bus route for airport access, with the objective of minimizing the total access time and used the artificial bee colony (ABC) approach to solve the model. He found under this method, the total travel time was less than the normal. Pratelli and Schoen [6] set up a mathematical model based on the balance of interests between passengers and waiting passengers, and optimized the location of demand response. They chose the appropriate solution for the model built by them, making the calculation speed of the model greatly improved, which laid the foundation for speeding up the response speed of passengers' reservation

system to the on-demand bus system. From the perspective of carpooling, Luca [7] studied the impact of charge based on distance and time on travel choice. The author found that travel cost, waiting time, gender, age and travel frequency were the most important factors that influenced travellers' choice of transportation modes. Liu [8] thought that distance charge can truly reflect passengers' travel costs, and played a regulatory role in the demand and supply of fares. Referring to the fare of taxis and regular buses, Liu [9] used the cost plus pricing method to determine the on-demand bus fare. The fare determined by this method was easier to be accepted.

The previous studies focused more on the influence factors, bus fare, travel time, travel cost, etc. They all got outstanding achievements. However, they all considered the problems separately and studied little on the passengers' waiting time [10–13]. This paper focuses on the passengers' waiting time and the bus profit and puts forward a mechanism of bus fare compensation and a mechanism of reservation cancelling. Based on the two mechanisms, on the case of fixed origin points and destination points, a model aiming at the maximum profit and the minimum total waiting time is established.

2 Model Establishment

2.1 Variables and Parameters Description

Parameters	Parameters description
N	Bus capacity
X_0	Initial bus fare
t	The time that the bus begins to accept an order
t_{i}	$\mathrm{i} = 0, 1, \ldots, \beta$, each expecting departure time
t_β	The latest time for departure
T	The time interval between the two expected departure times, which is called the valid request time
α	Fare discount
n_0	The number of the passengers reached between t and t_0
n_i	The number of the passengers reached between t_{i-1} and t_i
n'_{0j}	The number of the passengers reaching between t and t_0 and leaving between t_{j-1} and t_j, $\mathrm{j} \geq 0$
n'_{ij}	The number of the passengers reaching between t_{i-1} and t_i leaving between t_{j-1} and t_j, $\mathrm{i} < j$ and $\mathrm{i} \geq 1, \mathrm{j} \geq 2$
y	The cost of a bus running once
m	The cost of a bus running per kilometer

(continued)

(continued)

Parameters	Parameters description
L	Running mileage between the origin point and the destination point
m_0	The fixed cost of a bus running once, including shares of buses purchasing, personnel wages, etc.
λ	Exponential distribution parameter
f(x)	Exponential distribution density function
γ	The expected profit margin
ρ	The minimum guaranteed profit margin
t_{ip}	The Pth passenger arriving at $t_{i-1} \sim t_i$
$t_j - t_{ip}$	The Pth passenger arriving at $t_{i-1} \sim t_i$ when the bus leaves At t_j
T_{sumi}	The total waiting time of the passengers, which is the total waiting time of the passengers who are on the bus after departure (that is, the waiting time of the passengers who cancelled reservations is not taken into consideration)
$\overline{T_i}$	The average waiting time, which is the ratio of the total waiting time to the total number of passengers on the bus after the departure
z	The income of one running bus
z_1	The profit of one running bus
k	The times of departure delaying

2.2 Model Hypotheses

Some hypotheses were proposed during modelling:

(1) Passengers send out requests, choose the expected boarding time and ensure that they can get on the bus before departure;
(2) The running distance between any origin point and destination point in the road network is known;
(3) During each travel, the probability of passengers cancelling reservations obeys the exponential distribution;
(4) It is considered that the bus is running at a constant speed in the road network, not considering the influence of traffic conditions on the speed;
(5) The model is based on the operational characteristics of the mature on-demand transit service market;
(6) The waiting time of the passengers who cancelled the reservation during the waiting period is not included in the total waiting time;
(7) Only the benefits of single bus was taken into account;
(8) The passengers who send an order between t_i and t_{i+1} can only cancel the order after t_{i+1};
(9) When receiving N orders, the system will not accept the extra orders, and the extra orders will be assigned to the next bus.

2.3 Variables of the Model

α—the bus fare discount per extension of waiting time, $0 < \alpha < 1$;
T—the time interval between two expected departure times between t_{i-1} and t_i, which is called valid request time, $\mathrm{T} = t_i - t_{i-1}$.

When solving the model, to ensure α and T are not extremely big and extremely small, the values of α and T were given constraints.

$$\alpha_{min} \leq \alpha \leq \alpha_{max}$$
$$T_{min} \leq \mathrm{T} \leq T_{max}$$

$\alpha_{min}, \alpha_{max}, T_{min}, T_{max}$ are determined according to the spatial distance between the sites and the period of bus receiving the orders (Fig. 1).

At any time after the bus receiving orders, once the number of passengers is up to N or the delaying departure time reaches the threshold, the bus departs.

Every morning at a fixed time, the first bus begins to accept orders. In the meanwhile, the initial bus fare is X_0 and the expected departure time is t_0. At any time before t_0, once the reservation number reaches N, the bus can depart; otherwise, if after the first valid request time, T, namely t_0 moment, the total orders are still less than N, the departure time is pushed to $t_1 (= t_0 + \mathrm{T})$. At the moment, there are two choices for passengers: to cancel the order or continue to wait. For the passengers continuing to wait, their fare would be discounted to αX_0.

At any time during $t_0 \sim t_1$, once the reservation number is up to N, the bus can depart; otherwise, if after the second valid request time, T, namely t_1 moment, the total orders are still less than N, the departure time is pushed to $t_2 (= t_1 + \mathrm{T})$. At the moment, there are two choices for passengers: to cancel the order or continue to wait. For the passengers continuing to wait, the fare of the passengers arriving during $\mathrm{t} \sim t_0$ would be discounted to $\alpha^2 X_0$, and the fare of those arriving during $t_0 \sim t_1$ would be discounted to αX_0.

.....

At any time during $t_{\beta-2} \sim t_{\beta-1}$, once the reservation number is up to N, the bus can depart; otherwise, if after the second valid request time, T, namely $t_{\beta-1} (= t_{\beta-1} + \mathrm{T})$ moment, the total requests are still less than N, he departure time is pushed to t_β. At the moment, there are two choices for passengers: to cancel the order or continue to

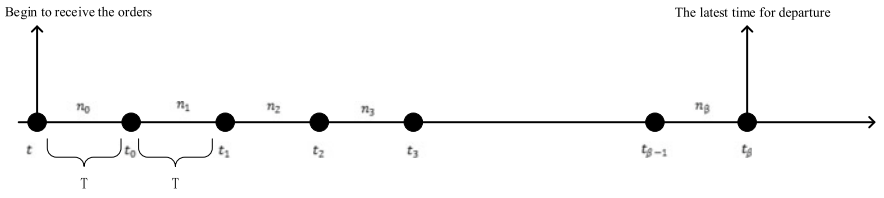


Fig. 1 Diagram of bus departure

wait. For the passengers continuing to wait, (1) for the passengers arriving during t $\sim$ t_0, the fare would be discounted to $\alpha^{\beta} X_0$, (2) for the those arriving during $t_0 \sim t_1$, the fare would be discounted to $\alpha^{\beta} X_0$, …, (β) for the those arriving during $t_{\beta-2} \sim t_{\beta-1}$, the fare would be discounted to αX_0.

At t_{β}, the bus departs regardless of the number of orders. And the fares of different passengers: (1) for the passengers arrive during t $\sim$ t_0, the fare is $\alpha^{\beta} X_0$, (2) for those arrive during $t_0 \sim t_1$, the fare is $\alpha^{\beta} X_0$, …, (β) for those arrive during $t_{\beta-2} \sim t_{\beta-1}$, the fare is αX_0, (β + 1) for those arrive during $t_{\beta-2} \sim t_{\beta-1}$, the fare is X_0.

2.4 The Cost of a Single Bus (Y) and Initial Ticket Price (X_0)

The cost of single bus, y, is determined by bus operation cost and the fixed cost. The fixed cost includes share equally of buses purchasing, personnel wages, etc.

$$\mathrm{y} = \mathrm{mL} + m_0 \tag{1}$$

where:

y—the cost of a single bus running
m—cost of unit mileage when bus running
L—running mileage of the origin point and the destination point
m_0—constant, consisting of buses purchasing share, personnel wages, etc.

An expected profit rate, γ, and a minimum guaranteed profit rate, ρ, are defined. The expected profit rate is the profit rate expected by the carrier. The minimum guaranteed profit rate is the profit rate that the carrier can get regardless of the fare discount, which is guaranteed.

The initial bus fare, X_0, is determined by the expected profit rate, the bus capacity and the cost of the bus. That is

$$X_0 = (1 + \gamma)\frac{(\mathrm{mL} + m_0)}{N} \tag{2}$$

According to the minimum guaranteed profit rate, ρ, there is following constraint

$$\alpha^k (1 + \gamma) \geq (1 + \rho) \tag{3}$$

2.5 The Relationship Between the Passengers' Waiting Time and the Waiting Behaviour

During the valid request time, T, there will be passengers who do not want to wait and choose to cancel the reservations and others who are willing to wait until the next expected departure time because of the fare discount.

The probability of a passenger not cancelling reservation halfway is assumed to obey exponential distribution. The shadow section in Fig. 2 represents the probability of passengers choosing to leave before t_i when the waiting time is T; the non-shaded part represents the probability of passengers choosing to wait until t_i, that is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4}$$

f(x) is the probability density function of exponential distribution, there is

$$\int_0^{+\infty} \lambda e^{-\lambda x} dx = 1 \tag{5}$$

where: λ is the parameter of the exponential distribution and its value can be determined by the investigation of the waiting passengers' psychology and be assumed.

From Fig. 2, in the case of invariant of the upper and lower bound of the integral (that is, when T is a fixed value), the bigger the λ, the smaller probability of passengers

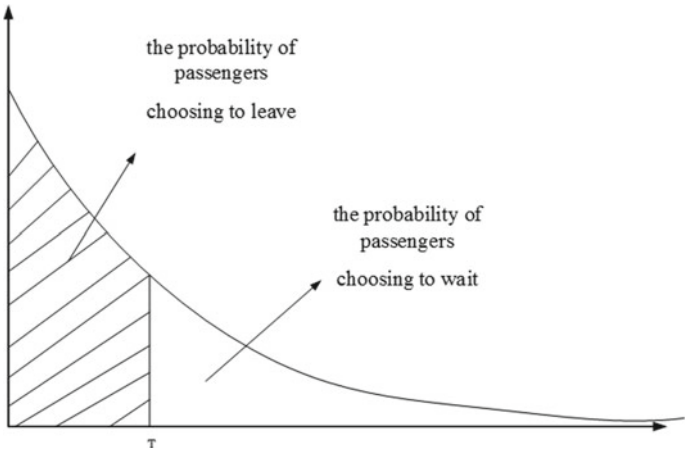


Fig. 2 Probability density of exponential distribution

leaving (the smaller the value of $\int_0^{+\infty} f(x)\mathrm{d}x$), the greater probability of passengers continuing to wait.

2.6 Single Bus Income Model

The single bus profit is the difference between the single bus income and the cost. The single bus income is equal to the sum of products of the number of passengers arriving at different moments and fares of different moments. According to the fare compensation mechanism, the bus fares are different for the passengers arriving at different time.

According to the relationship between the waiting time and the waiting behaviour and the fare compensation, the number of passengers at all moments can be obtained.

$$t_0\colon n_0$$

$$t_1\colon \left(n_0 - n'_{01}\right) + n_1 = \mathrm{floor}\left[n_0 \cdot \int_0^{\mathrm{T}} f(x)\mathrm{d}x\right] + n_1 \tag{6}$$

$$t_2\colon \left(n_0 - n'_{01} - n'_{02}\right) + \left(n_1 - n'_{12}\right) + n_2 = \mathrm{floor}\left[\left(n_0 - n'_{01}\right) \cdot \int_0^{T_2} f(x)\mathrm{d}x\right] + \mathrm{floor}\left[n_1 \cdot \int_0^{\mathrm{T}} f(x)\mathrm{d}x\right] + n_2 \tag{7}$$

$$\begin{aligned} t_3\colon & \left(n_0 - n'_{01} - n'_{02} - n'_{03}\right) + \left(n_1 - n'_{12} - n'_{13}\right) + \left(n_2 - n'_{23}\right) + n_3 \\ &= \mathrm{floor}\left[\left(n_0 - n'_{01} - n'_{02}\right) \cdot \int_0^{\mathrm{T}} f(x)\mathrm{d}x\right] + \mathrm{floor}\left[\left(n_1 - n'_{12}\right) \cdot \int_0^{\mathrm{T}} f(x)\mathrm{d}x\right] \\ &\quad + \mathrm{floor}\left[n_2 \cdot \int_0^{\mathrm{T}} f(x)\mathrm{d}x\right] + n_3 \end{aligned} \tag{8}$$

$$\cdots\cdots$$

$$\begin{aligned} t_k\colon & \left(n_0 - \sum_{i=1}^{k} n'_{0i}\right) + \left(n_1 - \sum_{i=2}^{k} n'_{1i}\right) \\ &\quad + \cdots + \left(n_{k-2} - \sum_{i=k-1}^{k} n'_{k-2i}\right) + \left(n_{k-1} - n'_{k-1k}\right) + n_k \\ &= \mathrm{floor}\left[\left(n_0 - \sum_{i=1}^{k-1} n'_{0i}\right) \cdot \int_0^{\mathrm{T}} f(x)\mathrm{d}x\right] \\ &\quad + \mathrm{floor}\left[\left(n_1 - \sum_{i=2}^{k-1} n'_{1i}\right) \cdot \int_0^{\mathrm{T}} f(x)\mathrm{d}x\right] \end{aligned}$$

$$+\cdots+\text{floor}\left[\left(n_{k-2}-n'_{k-2k-1}\right)\cdot\int_0^{\mathrm{T}} f(x)\mathrm{d}x\right]$$
$$+\text{floor}\left[n_{k-1}\cdot\int_0^{\mathrm{T}} f(x)\mathrm{d}x\right]+n_k \tag{9}$$

According to the fare compensation mechanism, for the passengers arriving at different time, the ticket price is different. At t_k, the ticket prices are as following:

passengers arriving at t $\sim$ t_0: $\alpha^k X_0$
passengers arriving at $t_0 \sim t_1$: $\alpha^{k-1} X_0$
passengers arriving at $t_1 \sim t_2$: $\alpha^{k-2} X_0$
passengers arriving at $t_{k-2} \sim t_{k-1}$: $\alpha^{k-1} X_0$
passengers arriving at $t_{k-1} \sim t_k$: αX_0

At t_k, the income of a single bus is

$$\begin{aligned}&\left(n_0-n'_{01}-n'_{02}-n'_{03}-\cdots-n'_{0k}\right)\alpha^k X_0\\&+\left(n_1-n'_{12}-n'_{13}-n'_{14}-\cdots-n'_{1k}\right)\alpha^{k-1}X_0\\&+\cdots+\left(n_{k-2}-n'_{k-2k-1}-n'_{k-2k}\right)\alpha^2 X_0\\&+\left(n_{k-1}-n'_{k-1k}\right)\alpha X_0+n_k X_0\end{aligned} \tag{10}$$

So, the single bus profit model is

$$\begin{aligned}z_1=&\left\{\text{floor}\left[\left(n_0-\sum_{i=1}^{k-1}n'_{0i}\right)\cdot\int_0^{\mathrm{T}} f(T)\mathrm{d}T\right]\cdot\alpha^k X_0\right.\\&+\text{floor}\left[\left(n_1-\sum_{i=2}^{k-1}n'_{1i}\right)\cdot\int_0^{\mathrm{T}} f(T)\mathrm{d}T\right]\\&\cdot\alpha^{k-1}X_0+\cdots+\text{floor}\left[\left(n_{k-2}-n'_{k-2k-1}\right)\cdot\int_0^{\mathrm{T}} f(T)\mathrm{d}T\right]\cdot\alpha^2 X_0\\&\left.+\text{floor}\left[n_{k-1}\cdot\int_0^{\mathrm{T}} f(T)\mathrm{d}T\right]\cdot\alpha X_0+n_k X_0\right\}-(\mathrm{mL}+m_0)\end{aligned} \tag{11}$$

2.7 The Model of Passenger's Waiting Time

The passenger's total waiting time, T_{sumi}, is the sum of the waiting time of passengers on the bus after departure (that is, the waiting time of passengers leaving halfway is not in consideration); the average waiting time, $\overline{T_i}$, is the ratio of the total waiting time and the number of passengers on the bus.

The total waiting time at t_k:

$$
\begin{aligned}
z_2 = & \left(n_0 - \sum_{i=1}^{k} n'_{0i}\right) \cdot t_k - \sum_{i=1}^{\left(n_0 - \sum_{j=1}^{k} n'_{0j}\right)} t_{0i} + \left(n_1 - \sum_{i=2}^{k} n'_{1i}\right) \cdot t_k \\
& - \sum_{i=1}^{\left(n_1 - \sum_{j=2}^{k} n'_{1j}\right)} t_{1i} + \left(n_2 - \sum_{i=3}^{k} n'_{2i}\right) \cdot t_k \\
& - \sum_{i=1}^{\left(n_2 - \sum_{j=3}^{k} n'_{2j}\right)} t_{2i} + \cdots + \left(n_{k-2} - \sum_{i=k-1}^{k} n'_{(k-2)i}\right) \cdot t_k \\
& - \sum_{i=1}^{\left(n_k - \sum_{j=k-1}^{k} n'_{(k-2)j}\right)} t_{(k-2)i} + \left(n_{k-1} - n'_{(k-1)k}\right) \cdot t_k \\
& - \sum_{i=1}^{\left(n_{k-1} - n'_{(k-1)k}\right)} t_{(k-1)i} + n_k \cdot t_k - \sum_{i=1}^{n_k} t_{ki}
\end{aligned} \tag{12}
$$

The average waiting time at t_k:

$$
\begin{aligned}
\overline{T_k} = & \left(\sum_{i=1}^{k} n_i - \sum_{i=1}^{k} n'_{0i} - \sum_{i=2}^{k} n'_{1i} - \sum_{i=3}^{k} n'_{2i} - \cdots - \sum_{i=k-1}^{k} n'_{(k-2)i} - n'_{(k-1)k} \right. \\
& - \sum_{i=1}^{\left(n_0 - \sum_{i=1}^{k} n'_{0i}\right)} t_{0i} - \sum_{i=1}^{\left(n_1 - \sum_{i=2}^{k} n'_{1i}\right)} t_{1i} - \sum_{i=1}^{\left(n_2 - \sum_{i=3}^{k} n'_{2i}\right)} t_{2i} - \cdots - \sum_{i=1}^{\left(n_k - \sum_{i=k-1}^{k} n'_{(k-2)i}\right)} t_{(k-2)i} \\
& \left. - \sum_{i=1}^{\left(n_{k-1} - n'_{(k-1)k}\right)} t_{(k-1)i} - \sum_{i=1}^{n_k} t_{ki} \right) \\
& / \left[\left(n_0 - \sum_{i=1}^{k} n'_{0i}\right) + \left(n_1 - \sum_{i=2}^{k} n'_{1i}\right) + \cdots + \left(n_{k-2} - \sum_{i=k-1}^{k} n'_{k-2i}\right) \right. \\
& \left. + \left(n_{k-1} - n'_{k-1k}\right) + n_k \right]
\end{aligned} \tag{13}
$$

Above all, a multi-objective optimization model is built:

$$
\begin{aligned}
max z_1 = & \left\{ \text{floor} \left[\left(n_0 - \sum_{i=1}^{k-1} n'_{0i} \right) \cdot \int_0^{\mathrm{T}} f(T) \mathrm{d}T \right] \cdot \alpha^k X_0 \right. \\
& \left. + \text{floor} \left[\left(n_1 - \sum_{i=2}^{k-1} n'_{1i} \right) \cdot \int_0^{\mathrm{T}} f(T) \mathrm{d}T \right] \cdot \alpha^{k-1} X_0 \right.
\end{aligned}
$$

$$+\cdots+\text{floor}\left[\left(n_{k-2}-n'_{k-2k-1}\right)\cdot\int_0^{\mathrm{T}} f(T)\mathrm{d}T\right]\cdot\alpha^2 X_0$$
$$+\ \text{floor}\left[n_{k-1}\cdot\int_0^{\mathrm{T}} f(T)\mathrm{d}T\right]\cdot\alpha X_0+n_k X_0\Bigg\}$$
$$-\ (\mathrm{mL}+m_0) \tag{14}$$

$$\begin{aligned}\min z_2=&\left(n_0-\sum_{i=1}^{k}n'_{0i}\right)\cdot t_k-\sum_{i=1}^{\left(n_0-\sum_{j=1}^{k}n'_{0j}\right)}t_{0i}+\left(n_1-\sum_{i=2}^{k}n'_{1i}\right)\cdot t_k\\&-\sum_{i=1}^{\left(n_1-\sum_{j=2}^{k}n'_{1j}\right)}t_{1i}+\left(n_2-\sum_{i=3}^{k}n'_{2i}\right)\cdot t_k-\sum_{i=1}^{\left(n_2-\sum_{j=3}^{k}n'_{2j}\right)}t_{2i}\\&+\cdots+\left(n_{k-2}-\sum_{i=k-1}^{k}n'_{(k-2)i}\right)\cdot t_k\\&-\sum_{i=1}^{\left(n_k-\sum_{j=k-1}^{k}n'_{(k-2)j}\right)}t_{(k-2)i}+\left(n_{k-1}-n'_{(k-1)k}\right)\cdot t_k\\&-\sum_{i=1}^{\left(n_{k-1}-n'_{(k-1)k}\right)}t_{(k-1)i}+n_k\cdot t_k-\sum_{i=1}^{n_k}t_{ki}\end{aligned} \tag{15}$$

S.T.

$$\left(n_0-\sum_{i=1}^{k}n'_{0i}\right)\geq 0 \tag{16}$$

$$\left(n_1-\sum_{i=2}^{k}n'_{1i}\right)\geq 0 \tag{17}$$

… … …

$$\left(n_{k-2}-\sum_{i=k-1}^{k}n'_{k-2i}\right)\geq 0 \tag{18}$$

$$\left(n_{k-1}-n'_{k-1k}\right)\geq 0 \tag{19}$$

$$\alpha^k(1+\gamma)\geq(1+\rho) \tag{20}$$

$$X_0>0 \tag{21}$$

$$n_i \geq 0 \tag{22}$$

$$\mathrm{T} = t_i - t_{i-1} \tag{23}$$

$$n'_{0j} \geq 0 \tag{24}$$

$$n'_{ij} \geq 0 \tag{25}$$

$$t_\beta = t + kT \tag{26}$$

$$\alpha_{min} \leq \alpha \leq \alpha_{max} \tag{27}$$

$$T_{min} \leq \mathrm{T} \leq T_{max} \tag{28}$$

3 Data Processing

The data of this paper came from the samples of DIDI's orders. The data of the period of 7:00:00–7:40:00 on December 22, 2016, in Beijing, was analyzed using the kernel density estimation based on the travel distance. The principle is as follow:

$$f_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K_h\left(\frac{x - x_i}{h}\right) \tag{29}$$

$K_h(x)$ is the kernel function [14]. There are many kinds of kernel functions, such as Gaussian, uniform, triangular, Biweight, triweight, Epanechnikov, normal and so on. h is a smooth parameter, called bandwidth (bandwidth), also called a window. This paper uses the Gaussian kernel function and the default optimal bandwidth to estimate the kernel density.

ArcGIS was used to show the distribution of the starting points and the ending points. It can be seen that most of the starting points and ending points are concentrated in the central city and a few of them are located in Tongzhou, airports and other places far away from the central city (Figs. 3 and 4).

The model established in this paper is based on the following characteristics: the starting points and the ending points are within the same small radius. They occurred in the areas where trips are not too dense. Therefore, according to the characteristics of the model, the data whose travel distance was larger than 25,000 m were selected and 79 sets of data were obtained. Then, the 79 sets of data were used to test the model.

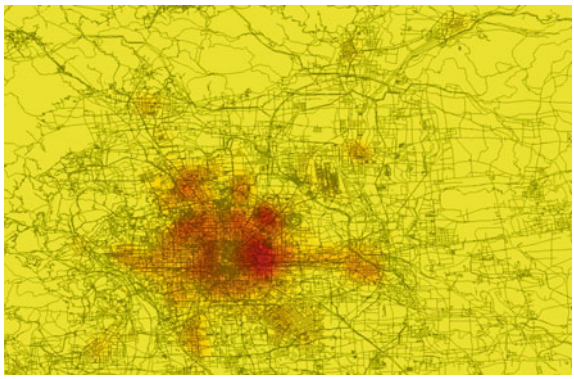


Fig. 3 The distribution of starting points

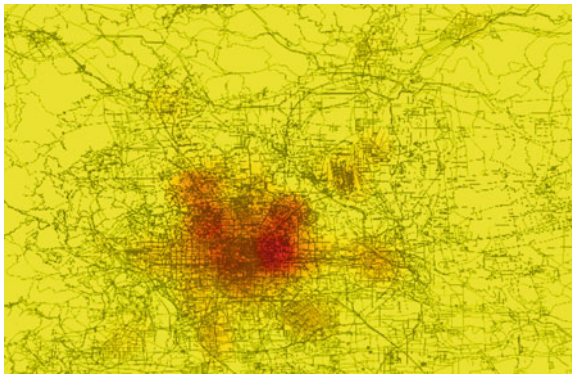


Fig. 4 The distribution of ending points

4 Model Solution

We choose Yizhuang Development Zone and Tongzhou District in Beijing as OD points. The parameters assignment are shown in Table 1.

From the Tongji Nan Road subway station in Yizhuang Development Zone to the Tongzhou District government, through the East sixth ring road, the travel distance is about 35.7 km, so the running mileage, L, is assigned as 36 km. According to the above assignment parameters, it can be calculated that the single bus cost is 174 yuan. We select the minibus whose capacity is 20 and make the expected profit margin 1 and minimum guaranteed profit margin 0.5. From the above, the original bus fare is 17.4 yuan. The operating bus begins to accept requests at 7:00:00, and the latest time for departure is to be 40 min after accepting orders, that is $k = 3$.

According to the parameters assignment, LINGO was employed to solve the model. LINGO is the best choice for optimization model. For multi-objective optimization, the common methods include main objective, linear weighted summation, exponential weighted product, etc. In this paper, the linear weighted summation

Table 1 Parameters assignment

Parameters	Parameter values
N	20
X_0	17.4
t	7:00:00
y	174
m	4
L	36
m_0	30
λ	600
γ	1
ρ	0.5
T_{min}	300
T_{max}	600
α_{min}	0.8
α_{max}	0.95
k	3

method is used. ω_1 is the weight of bus profit function, and ω_2 is the weight of waiting time function. The following results are obtained by the different values of ω_1 and ω_2.

From Table 2, with the two weights changing, the optimal solutions of the model change. With the increase of the weight of the first objective, the bus profit gradually increases and the total waiting time decreases. Meanwhile, valid requests time gradually increases with ω_1 increasing, which means the waiting time of passengers choosing to wait increases, conforming to the actual condition. However, the optimal solution of fare discount, α, remains unchanged, which means relevant parameters should be adjusted.

Transfer times, sites number, travel time and travel cost of different travel methods are compared in Table 3.

From Yizhuang to Tongzhou, whatever by bus or subway, it is a travel of long-time and many-transfers. If choosing the taxi, the travel cost is too high to afford for most passengers. Therefore, the on-demand bus may be a better alternative.

Tab. 2 The optimal solutions of the model with different weights

ω_1	ω_2	α	T(s)	maxz_1(yuan)	min$z_2(s)$	$\overline{T}(min)$	Departure time
0.7	0.3	0.95	305.0895	96.1183	11671.8	9.73	7:20:20
0.75	0.25	0.95	419.0728	133.037	20790.42	17.33	7:27:56
0.8	0.2	0.95	466.5133	140.037	24585.07	20.49	7:31:06

Tab 3 The comparison of different transportation

Transportation	Transfer times	Sites number	Travel time	Travel cost
Subway	1	28	1 h and 28 min	9 yuan
Traditional bus	1	27	1 h and 40 min	7 yuan
Taxi	–	–	50 min	93 yuan
On-demand bus	–	–	1 h and 10 min	17.4 yuan

5 Conclusion

In this paper, the mechanism of bus fare compensation and the mechanism of reservation cancelling are put forward. Based on the two mechanisms, the on-demand bus model is built and tested by the data from Didi. According to the results, the travel time is shorter than traditional bus and the cost is lower than the taxi.

The paper explores an on-demand bus considering the passengers' waiting time. The on-demand bus can make full use of public transportation resources and improve the bus service level to a certain extent. It provides a theoretical basis for the establishment of the on-demand bus systems.

Acknowledgements This paper study was sponsored by DIDI Chuxing, Key Laboratory of Road and Traffic Engineering, Ministry of education, Tongji University (No. TJDDZHCX012), the Special Funds For Basic Research Business Fees of Central Universities of Chang'an University (No. 300102218409), and Technical Innovation Funding of China Academy of Urban Planning & Design, Traffic congestion mechanism analysis based on data mining technology (No. C-201728). The author is grateful to the teacher who guided her a lot and the classmates who worked with her. The author is also grateful to the DIDI Chuxing for supporting data. The remaining shortcomings are the author's alone.

References

1. Liu Y (2015) Study on the planning and development evaluation of urban custom bus line. Xi'an University of Architecture and Technology
2. Guo X (2016) Research on real-time on-demand bus system based. Beijing Jiaotong University, Beijing
3. Li B (2013) Research on on-demand bus. Chang'an University
4. Huang D, Liu Z, Liu P, Chen J (2016) Optimal transit fare and service frequency of a nonlinear origin-destination based fare structure. Transport Res Part E
5. Jinux C (2017) Design of suburban bus route for airport access. Transportmetrica A Transp Sci
6. Schoen F, Pratellil A (2001) A mathematical programming model for the bus deviation route problem. J Oper Res Soc 5(52):494–502
7. Luca S, Di Pace R (2015) Modelling users' behaviour in inter-urban carsharing program: a stated preference approach. Transp Res Part Policy Prac 71:59–76
8. Liu T, Ceder AA (2015) Analysis of a new public-transport-service concept: customized bus in China. Transp Policy 39:63–76
9. Liu X (2015) Study on urban custom bus line design. Changsha Univ Sci Technol

10. Chen PW, Nie YM (2017) Analysis of an idealized system of demand adaptive paired-line hybrid transit. Transport Res Series B Methodol
11. Wong KI, Bell MGH (2006) Solution of the dial-a-ride problem with multi-dimensional capacity constraints. Trans Operational Res
12. Han B, Fan J, Wu H (2016) Analysis and planning of community bus in Suzhou. In: The collection of China annual conference on urban traffic planning
13. Wu L (2014) A flexible route bus scheduling method considering passengers' waiting behavior. Dalian University of Technology
14. Lin M, Zhen F (2015) Research on resident activity space of Nanjing city based on nuclear density analysis, annual meeting of China's urban planning

Research on the Satisfaction Degree of Rookie Station Based on Centrality Analysis—Taking Shenzhen University as an Example

Hui Yin and Liang Zou

Abstract In recent years, more and more colleges and universities join the rookie station, however, the users' satisfaction with the service is not very high. This paper takes rookie station user satisfaction degree as the cut-in point, using the centrality analysis method of social network analysis, taking the actual investigation and the network questionnaire as the data obtaining way, exploring the network relationship between the users and the satisfaction factors of the rookie station, and then provides the improvement plans for the service quality of the rookie station. And taking the rookie station of Shenzhen University as an example, through 100 effective questionnaires, using UCINET software, the centrality of the factors are analyzed. The results show that campus users pay more attention to the tally speed and spatial layout, but they pay less attention to the shipping charge standard and pick-up mode.

Keywords College rookie post · Centrality analysis · Satisfaction degree

1 Introduction

In recent years, high school express rapid development, according to the first "Campus Express Industry Development Report (2016)" shows that the whole year 2015, the national universities received more than 720 million packages, accounting for the total number of 6% [1]. With the increasing number of parcels in colleges and universities, the terminal distribution link becomes more and more important. Since 2013, Cainiao network began to establish community-oriented and university logistics service platform. Up to now, the country has set up more than 2500 college rookie stations. At the same time, the rapid development of the market of college rookie station is accompanied by many problems which need to be solved urgently, such as bad service attitude, long queue time for pick-up, parcels litter and damage and so on [2]. According to the "2016 Express Market Monitoring Report" on the Express service complaints can be seen, consumer complaints express enterprise's

H. Yin (✉) · L. Zou
College of Civil Engineering, Shenzhen University, Shenzhen 518060, China
e-mail: 570096307@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_36

core problems are mainly delivery services, 99,144 pieces, accounting for the total number of complaints 37.7% [3]. Among them, the proportion of complaints in colleges and universities is as high as 9%, which is more than the proportion of parcels 6%, which further illustrates that it is imminent to upgrade the level of express service in colleges and universities, especially the service level of rookie station.

At present, the research on the rookie station is mainly based on domestic scholars. Zhe Zhang through constructing the distribution satisfaction model, the paper sums up the distribution speed of goods, the distribution quality of goods and the quality of service, which affects the satisfaction of consumers to distribution [4]. Rushi Tan through the analysis of questionnaires and GIS, SPSS software, taking Nanjing City rookie station as an example, probed into the relationship between spatial characteristics of residents' behavior and spatial distribution of urban express and provided references and suggestions for its scientific distribution [5]. Silei Chen probed into the express operation mode of colleges and universities, combining with the present situation of many problems of express service in colleges and universities, and proposed to improve the service level and other suggestions [6]. To sum up, we can find that the current research is mainly focused on the operating mode, spatial layout, development direction and other macro-level of the college rookie station, but the micro-level research on its own environmental conditions, but the potential relationship between service factors and users is very small.

Social network analysis is an accurate quantitative analysis of the relationship between human beings and people and objects, which can reveal the intrinsic structure of the relationship and is widely used in the construction of the relationship theory and the empirical test [7]. Therefore, this paper will use the social network analysis of the centrality analysis of the College Rookie station user satisfaction factors of the impact of in-depth, comprehensive analysis, further define the rookie station each service quality respectively to the users' satisfaction influence, for the rookie station Service quality to provide improvement plans.

2 Construction of Social Network of Satisfaction Degree of College Rookie Station

Social Network refers to the collection of social actors and their relationships, that is, a social network is a collection of connections between multiple points and points, emphasizing that each of the actors has a more or less relationship with other actors [8].

2.1 Construction of Social Network of Satisfaction Degree of College Rookie Station

2.1.1 Determination of Network Nodes

Before the social network analysis, first of all to determine the network nodes, according to the different network types, network nodes are also diverse. In this paper, different users and their attention to the satisfaction of the factors as a node, through their connection to build the corresponding network.

(1) Social actors' nodes

The social actors in this paper are the students of the college freshmen, namely the Shenzhen University which is surveyed.

(2) Influence factor nodes

In this paper, according to the relevant references and the actual situation of college rookie station, puts forward the 14 influence factors that users may pay attention to, such as geographical location, business hours, spatial layout, goods placement, pick-up time, pick-up mode, service attitude, express delivery fee standards, pick-up fee standard, users' pick-up order, tally speed, package security, the collection of complex courier procedures, the users' personal information protection. The first step, through the questionnaire survey to the users' actual attention degree carries on the analysis, the results show that the attention rate of users' pick-up order, the geographical location, the users' personal information protection is low, only 0.12%, 0.1%, 0.8% and 0.3% respectively. So this paper chooses the 10 factors that the attention rate of above 8% that affect user satisfaction as the network influence factors nodes, they are the shipping charge standard (f1), pick-up line Time (f2), pick-up Mode (f3), goods placement (f4), Business hours (f5), Spatial Layout (f6), service attitude (f7), Pickup fee (f8), parcel Safety (f9), tally Speed (f10). The specific rate of attention is shown in Fig. 1.

2.1.2 The Determination of Social Network Relationship

After the social network node is determined, the next is to determine the edge of the social network, that is, to determine the connection between the social network nodes. Because this article is to affect the University rookie station user satisfaction research, it belongs to the 2-model network relationship.

The 2-mode relationship describes the relationship between a group of social actors and many events (indicators, events, affiliations), also known as the affiliation data [9]. The 2-mode relationship in this paper is mainly about the relationship between the user and the satisfaction degree of the rookie post in the university.

$$X_{ij} = \begin{cases} 0, \ (Actor\ i\ don't\ pay\ attention\ to\ factor\ j) \\ 1, \ (Actor\ i\ pay\ attention\ to\ factor\ j) \end{cases} \tag{1}$$

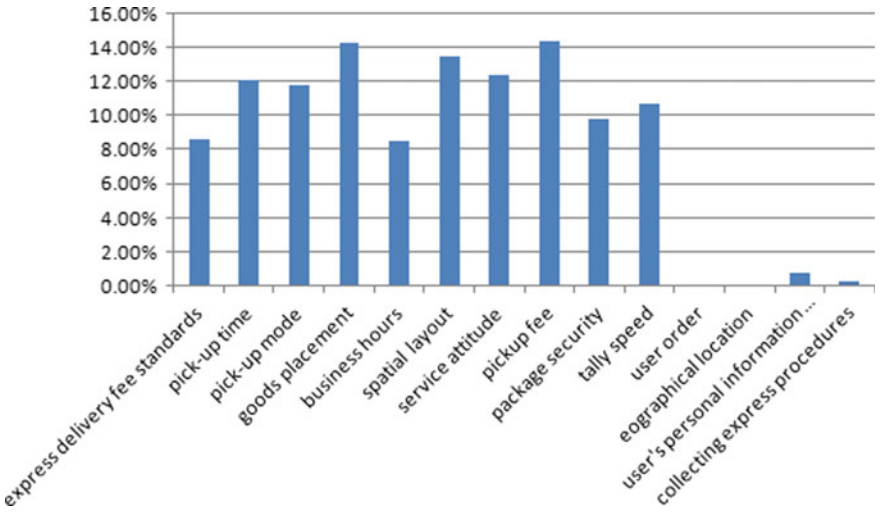


Fig. 1 Statistical table of factors affecting user satisfaction

($i = 1, \cdots, M$, $j = 1, \cdots, N$, M is the total number of actors, N is the total number of factors).

3 The Method of Centrality Analysis

In this section, we mainly use the centrality analysis method to analyze the factors that affect the user satisfaction degree. The centrality analysis in the 2-mode network is the quantitative index of the node "right" in the network graph, which reflects the importance degree of the node in the network, including the relative degree centrality, closeness centrality and betweenness centrality. This paper mainly analyzes the centrality of the factors in the 2-mode network.

(1) Relative degree centrality

In the bipartite "2-mode" graphs, the degree centrality of a factor is the number of actors owned by the factor. Standardize the centrality indicator, that is to say, the number of actors involved in one factor divided by the total number of actors then get the relative degree centrality of factor j, and the calculated formula is:

$$r_j = \frac{\sum_{i=1}^{M} X_{ij}}{M}, (j = 1, \cdots, N) \tag{2}$$

(2) Closeness centrality

In the bipartite "2-mode" graphs, the factors are only associated with the actors, therefore all the pathways from the factors must first pass through the actors involved

in the event. In a affiliate network, the closeness centrality c_j of a factor j is a function of the shortest distance from the actors involved in the factor to the other actors and factors, as follows:

$$C_j = \left[1 + \frac{\sum_{i=1}^{M} X_{ij} \left(\sum_{k_1=1}^{M} DM(i, k_1) + \sum_{k_2=1}^{N} DN(i, k_2) \right)}{M + N - 1} \right]^{-1} \tag{3}$$

$DM(i, k_1)$ represents the shortest distance a user i to another user k_1, $DN(i, k_2)$ represents the shortest distance a user i to another user k_2.

(3) Betweenness centrality

The betweenness centrality is concerned with the extent to which a factor resides in the middle of the network, taking into account all shortcuts at the time of calculation. In the affiliate network, the connection between the factors is to be done by the actors, so the actors are always in the shortcut between the factors. Therefore, the relative betweenness centrality b_j of factor j is:

$$b_j = \frac{\sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \left(X_{i_1 j} X_{i_2 j} \frac{1}{\sum_{k=1}^{N} X_{i_1 k} X_{i_2 k}} \right)}{2} \quad j = 1, \cdots, N \tag{4}$$

4 Example Data Collation and Centrality Analysis

Taking the students of Shenzhen University as the object of investigation. The questionnaire design was completed by using the methods of consulting the literature, the experts' opinions and so on, and the questionnaire was taken as the data collection method by the network questionnaire and take the form of a random questionnaire to the students within the whole school to start a questionnaire. A total of 121 questionnaires were collected, with 100 effective questionnaires and 82.6% effective recoveries. The overall status of the questionnaire is shown in Fig. 2.

As the subject of the survey for students, here with U to represent undergraduates, with G to represent graduate students, such as: UOW on behalf of undergraduate girl in the first year, with GTM on behalf of the second-year graduate boy.

4.1 User-Satisfaction Influence Factor 2-Model Data

The user to the rookie station satisfaction factors of the statistics, all users are concerned about factors in this factor is a column of 1, the factors that do not pay attention

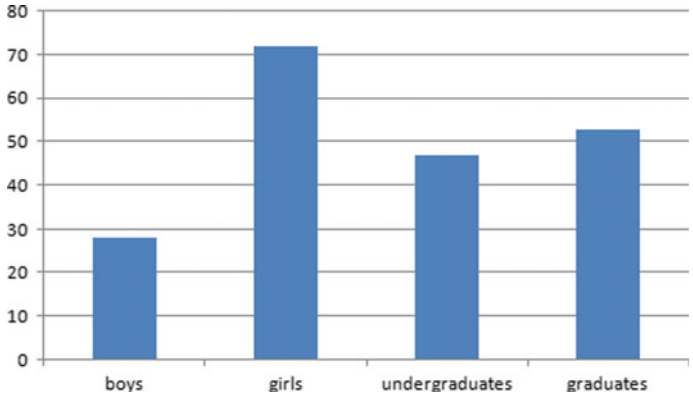


Fig. 2 Total statistics of questionnaires

to 0, get a user-satisfaction influence factor 2-Model network table, the following specific gives 5 users-10 satisfaction factors of the data, as shown in Table 1.

According to the statistic data of 2-model relationship, we can analyze the social network of the users and the satisfaction factors of the rookie station.

4.2 A Centrality Analysis of the Factors Influencing Satisfaction Degree

In the Ucinet software, the above surveyed users and their impact on the rookie station satisfaction factors to do a central analysis, the results are shown in Table 2

(1) Analysis of core factors

From the above table, it can be seen that the relative degree centrality, closeness centrality, betweenness centrality of f10 and f6 are higher in the 10 factors that affect the user satisfaction, which shows that the tally Speed and spatial layout of the goods are in the core of the network event, which is the focus of the users. This is mainly because they directly determine the user's terminal service experience, in which the tally Speed determines the timeliness of courier tracking information and the timeliness of delivery; the spatial layout determines the speed of pick-up and the overall intuitive feeling of the user for the rookie station.

(2) Analysis of marginal factors

The relative degree centrality, closeness centrality, betweenness centrality of f1, f3 are lower in the 10 factors, which indicates that the three factors of the shipping charge standard, the pick-up mode and Pickup fee standard are on the edge of the network event, and the users seldom pay attention to these factors. Through the field investigation and research, 99% of the students are to pick up pieces, and only 1% of

Table 1 User-Satisfaction factor 2-modulus datasheet (5)

Factor										
User	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10
GTW1	0	0	0	1	0	1	1	0	1	1
GTW2	0	0	0	1	0	0	1	1	0	1
GTW3	0	1	0	1	1	0	0	0	0	1
GTW4	0	1	0	0	0	1	0	0	0	1
GTW5	0	0	0	1	0	1	0	0	0	1

Table 2 Results of central analysis of factors

Centrality factor	Relative degree centrality	Closeness centrality	Betweenness centrality
f1	0.050	0.364	0.002
f2	0.270	0.444	0.070
f3	0.180	0.418	0.040
f4	0.240	0.434	0.042
f5	0.390	0.492	0.147
f6	0.570	0.573	0.313
f7	0.280	0.450	0.081
f8	0.250	0.440	0.067
f9	0.380	0.488	0.144
f10	0.520	0.551	0.248

the students are to mail items, so we basically do not pay attention to the cost of the shipping charge standard; users are also very little concerned about pick-up mode (f3). The school is different from the surrounding rookie station about pick-up way. The school's rookie station is the students to find their own pieces of goods code, and outside the rookie station is the staff to find their own pieces, because the school's rookie station space layout is larger, can accommodate a large number of students. On the one hand, this method of pick-up to the staff to reduce a lot of work, on the other hand, save the students queue, shorten the pickup time.

5 Conclusion

On the basis of the analysis results in the fifth section, this paper puts forward the opinion of improving the customer satisfaction of the freshmen in colleges and universities: The rookie station should carry on the difference management. Focusing on key factors such as speed and spatial layout. First of all, the staff should be trained in professional training, especially the tally business, improve the efficiency of the goods; secondly, the overall space layout should be further optimized to avoid the occurrence of narrow and disorderly space layout, as far as possible to achieve subregional management. However, the marginal factors, such as the shipping charge standard, pick-up mode and pick-up fee standard, rookie station can gradually reduce the investment of human and material resources.

Acknowledgements Thanks for the project fund: the sponsorship and support of the 2017 Shenzhen Science and Technology Project (JCYJ20170818142947240), I would like to express my deep appreciation.

References

1. Ali Research Institute (2016) Campus express industry development report
2. Huimin Du (2017) Analysis on current situation and development countermeasures of rookie station. Natl Circ Economy 28:14–15
3. People's Republic of China Post office (2017) 2016 annual express market monitoring report
4. ZZhang Z (2010) A study on the satisfaction evaluation system of logistics distribution service for users of net purchase. Logistics Technol 11:27–28
5. Rushi T, Xu Y, Dong C, Liu L (2016) A study on the behavior space of urban residents' express self-action. World Geography Stud 05:111–120
6. Chen S, Jiaoling LI, Yang Y (2012) A study on service quality of express enterprises under the operation mode. In: China conference, pp 87–91
7. Hao Z (2009) Blog community based on social network analysis. Shanghai Jiao Tong University, Shanghai, pp 10–18
8. Boykin PO, Roychowdhury VP (2005) Leveraging social networks to fight spam. IEEE Comput (04):61–68
9. Liu J (2004) Introduction to social network analysis. Social Sciences Literature Press

Evaluation Method of Drivers Vision Impacts from Green Belts of Arterial

Yu-gang Sheng, Wan-lu Song and Jian-xiao Ma

Abstract As an important part of the cross section of road, the green belt of urban road has heavy impacts on the drivers' visual environment. Green belt directly relates to traffic safety. But the research on this area is only carried out from the aspects of color and monotonicity. Few research can be found which is about impacts indexes. In this article, the impacts model of green belt is established. The green belts impact value of the different level and different section is obtained by expert method. In order to validate the model, an eye movement instrument (SmartEye) is adopted. The data of impacts indexes can be obtained in SmartEye. Then, ErgoLAB package is used to identify the drivers' gaze points during the course of the experiment and to generate the view distribution map of the gaze area. After this, the experiment in driving simulator dome under variable traffic flow condition is carried out to modify the model built above. At last, a recommended rating grade standard is presented.

Keywords Traffic safety · Green belts of urban arterial · SBE · Impact model · Eye tracker · Driving simulator

1 Introduction

The number of urban traffic accidents has risen in few years, as the proportion of motor vehicle drivers is up to 70%. As the driver of driving behavior, 90% of driving information is obtained through visual perception, so the importance of visual environment is elf-evident. The urban road green belt is an important part of the cross-section of the city, and it has a direct effect on the visual environment of the driver. With the acceleration of urbanization, more intensive urban road space has been created, which has changed the original road greening pattern, and new road structure has given rise to new green forms. While the road greening plants highlight the landscape and construct the image of the city, the influence of the road traffic

Y. Sheng (✉) · W. Song · J. Ma
College of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing 210037, China
e-mail: shengyugang@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_37

safety should also be considered [1]. So far, the study of the urban road green belt has been studied in terms of color and monotony, and it lacks quantitative research on other indicators of green belt [2].

During driving, the road green belts produce a corresponding stimulus to the driver's vision, and if the stimulus is too strong, it can cause too much of the driver to pay attention to the green belts, thus causing an in concentration of attention. Otherwise, drivers will feel bored and increase visual fatigue [3]. Green belts on the degree of stimulation driver, will affect the driver's driving attention, thus affecting the road emergencies and other security hidden trouble judging the rationality and reflection of the time. If the driver's fixation point is not reasonable and attention is not focused, it greatly increases the probability of road traffic accidents and affects road traffic safety.

Research work starts from the urban arterial green belts, SBE (Scenic Beauty Estimation Method) method is selected to evaluate the visual effects of the selected roads. And using SPSS mathematical statistical analysis software to establish the visual influence degree model of the green belt index [4]. In order to verify the reliability of the model, it is necessary to use eye tracker to record the fixation point distribution of the driver in the green belt's AOI (Area of the Interest Region). Using the driving simulator to study in the case of different traffic patterns where the characteristics of urban arterial green belts influence vision of the driver. On this basis, build a green belt scale model of different speeds. Finally, based on the expert scoring method, the evaluation level of the influence degree of green belt is established. For normal or poor visual effects, the scoring interval should be appropriate to the visual effect. Select the specific level of each factor reasonably, combine the evaluation model, and make the green belt score fall into the zone, give the driver a comfortable aesthetic feeling and proper alertness and excitement to ensure the stability of the driving [5].

2 Establish the Impacts Model of Green Belt

2.1 Selection of Typical Green Belts by LCJ Method

The LCJ method is the dual comparison method [6], collects the green belt photograph which needs to appraise, through the photograph 2 and 2 comparison, lets the evaluators pick out the green belt more common one, so repeatedly carries on the comparison and the selection, makes each picture to have with any other comparison opportunity. Finally, the test path is determined according to the chosen probability.

(1) Phase of Road Cross Section Selection

Prepare 70 pictures of road landscape, according to the road plate in the following classification comparison, the final selection of urban roads: Beijing East Road, Jinxianghe Road, Zhongshan North Road, Taiping North Road, the new model Road,

Table 1 Classification of road cross section

Form	Representative sections
One plates two belt	Lhasa Road, Sheng Zhou Road, Hunan Road
Two plates three belts	Jinxianghe Road, Xian Lin Road
Three plate four belts	Changjiang Road, Zhongshan North Road, Hongwu North Road, Beijing East Road, Taiping North Road
Four plates five belts	Xinmofan Road, Xuezi Road, Huitong Road, Xuehai Road

Hongwu North Road, the Changjiang Road, a total of seven roads as a test road (Table 1).

(2) Road Data collection phase

Driving vehicles at the speed of 40 km/h on the selected test road smooth, from the driver's perspective shooting, after filming through video clips to remove bumpy road, select a smooth section video as a follow-up test material.

2.2 Index Selection of Green Belt by Delphi Method

The selection of green belt indicators is determined by the Delphi method. Experts evaluate the impact of various indicators of the seven main road green belts, including Beijing East Road. Experts consider the characteristics of different tree species based on their own rich knowledge and life experience. The driving angle of view and the alternation of seasons are used to rank the different indicators of seven roads. The scoring criteria are 100 points, 75 points, 50 points, 25 points, 0 points. The greater the visual impact, the higher the score. The score table is shown in Table 2. The major expert opinions are comprehensively analyzed to obtain the degree of concentration of different indicators, the coefficient of variation, and the final determination of the indicators of greater concentration and smaller coefficient of variation [7] (Table 3)

(1) Concentration

$$M_j = \frac{1}{m_j} \sum_{i=1}^{m} C_{ij} \tag{1}$$

where: M represents the concentration of expert opinions, its size determines the importance of the index. m represents the number of experts participating in the JTH index score. C represents the score of the index. Subscript I is used to distinguish between experts, and subscript j is used to distinguish indicators.

Table 2 Eigenvalue in Delphi evaluation method

Name	Unit	Evaluation objects
The degree of visual impact refers to the degree of visual attraction of the greening on both sides of the road to the driver, which can be measured by the length of fixation time on the fixation point		
Evaluation's index	Evaluation's indicator description	Evaluation
1. Height	The distance from the ground up to the top of the tree, of which green ups and downs also belong to the category of height	
2. Tree form	Refers to the spatial structure of trees	
3. Sky ratio	Refers the effective sky area occupies the total area of the driver's field of view	
4. Trunk visibility	Refers the sensitivity of a driver to a tree trunk during driving	
5. Crown width	Refers the average width of the north or south of the tree	
6. Spacing	Refers the distance between two adjacent trees	

Table 3 Delphi method score characteristic value

	Height	Crown width	Tree form	Sky ratio	Trunk visibility	Spacing
Concentration	90	95	85	45	70	15
Coefficient of variation	0.15	0.12	0.16	0.25	0.16	1.49

(2) Coefficient of variation

$$\sigma_j = \frac{\sqrt{\sum_{i=1}^{m_j} (C_{ij} - M_j)^2}}{m_j} \tag{2}$$

$$V_j = \frac{\sigma_j}{M_j} \tag{3}$$

where: σ is the standard deviation. V represents the variation coefficient, and the variation coefficient mainly reflects the degree of coordination of expert opinions, it is an important index to evaluate relative fluctuation. The smaller the coefficient of variation, the more concentrated the opinions of the experts.

After testing, the eigenvalues of all indexes are arranged in numerical order of concentration. Experts believe that seasonal alternations lead to uncontrollable changes in the color of the leaves of the street trees and inconspicuous changes in the distance between the driver's perspective trees. The SBE method experiment was carried out

Table 4 Test score sheet

Age		Gender		Driving age	
The green belt index	Height	Crown width	Tree form	Sky ratio	Trunk visibility
Degree of visual influence					

to preserve the height, crown width, tree shape, sky ratio and trunk visibility of the green belt [8].

2.3 Testing Program Design of Visual Impacts Based on SBE

In the evaluation, the SBE method uses photographs or slides as the media for judgment. According to the evaluation criteria, the judges rate the green belt indicators of each group of photos according to the degree of visual distraction.

This experiment randomly selected 50 testers to conduct experiments. The experimental method was to allow the testers to watch the experiment video to score the visual impact of each indicator. The scoring standard adopted a 10-point system, allowing the testers to watch the road video. Select the three sections with the best visual effects and mark them (Table 4).

2.4 Analysis on Experiment Data

The SBE method scores all evaluation samples according to the influence degree, and then regroups all the score results according to different greening indexes. To study the quantitative relationship between index scores and their eigenvalues, the metrics of each index are measured as follows:

The height indicator of the green belt can be measured with the aid of a laser range finder. The height of the tree is measured using the triangle pylon principle (pitch, elevation angle and horizontal distance). The measurement result is expressed in meters.

The crown index of the green belt can be measured by the crown projection method. Assume that the projection of the crown on the ground is elliptical, and the major axis and the minor axis of the ellipse are measured, taking the mean value, and the measurement result is measured in meters.

The green belt sky ratio indicator can be recorded with the camera at the driver's first viewpoint. After the video is completed, the ratio of the sky area to the entire screen area is measured for each frame using the grid method to obtain the average value.

The measure of the visibility of the trunk of the green belt is different. Because the distance between the driver and the tree body is different, the visibility of the trunk is different from that of the driver. The closer the distance, the larger the tree appears, and vice versa. Therefore, the trunk visibility is determined as the product of the driver's visual distance and the visible trunk area at the driver's perspective, in units of: $100m^3$.

The determination of the tree-shaped index of the green belt is determined according to the shape of the outer contour of the tree. The tree of the trunk roadway can be roughly divided into a spire, a sphere, and an umbrella. The same principle as the determination of the visibility of the trunk is determined by the product of the driver's visual distance and the area of the visible part above the trunk at the driver's perspective, in units of: $100m^3$.

Data statistics were collected to obtain the feature values and scores of different road indicators (Table 5).

2.5 Drivers Visual Impacts Modelling

The process of building a model can be divided into the following three steps:

(1) Analyze experimental data and remove abnormal data

This evaluation model is based on the fact that the characteristic index of the plant index is an independent variable and the score value is established on the basis of the dependent variable. In order to enhance the usability and accuracy of the evaluation results, the abnormal data was removed using the Grubbs test [9]. Calculate the average score x and standard deviation S of the same indicator for different drivers. Calculate the G value and check the Grubbs test table (Table 6) based on the number of measurements and confidence requirements. Compare G calculation with G table, if G calculate > G table, discard, otherwise keep it.

$$S = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{4}$$

where: S is the standard deviation value; n is the total number of drivers.

$$Gc = \frac{X_n - \overline{X}}{S} \tag{5}$$

where: X is the score, n is the number of grader.

(2) Make reasonable assumptions and analyze the internal quantitative relationship

There are obvious differences in the heights of different green belts. Drivers will reduce their visual acuity as the vehicle speed increases. Therefore, when the tree

Table 5 Single indicator factor score

Road indicators	Species	Height (m)	Score	Crown width (m)	Score	Tree	Score	Sky ratio (%)	Score	Trunk visibility	Score
Beijing East Road	Cedar	14	4.34	7.5	6.11	3.01	5.41	28	2.36	1.3	1.41
Jinxiang River Road	Cedar	28	5.41	6.8	4.90	2.65	4.77	8.50	2.48	3.1	3.28
Hongwu North	Citron	8	3.45	8	2.97	1.7	3.41	18	2.42	1.8	1.76
Taiping North	Cedar	24	5.39	5.9	2.82	1.54	2.96	20	2.44	1.38	1.41
New Model Road	Citron	5.7	2.93	4.9	3.18	1.64	2.96	25	2.36	0.61	0.71
Changjiang Road	Indus	20	5.36	16	6.82	4.01	7.21	1	2.56	5.68	5.72
Zhongshan North Road	Indus	15	4.50	10	5.68	2.81	5.25	5	2.58	3.73	3.82

Table 6 Grubbbs inspection table

n / p	0.95	0.99	n / p	0.95	0.99
3	1.135	1.155	11	2.475	2.785
4	1.463	1.492	12	2.504	2.821
5	1.672	1.749	13	2.532	2.854
6	1.822	1.944	14	2.557	2.884
7	1.938	2.097	15	2.58	2.912
8	2.032	2.231	16	2.603	2.939
9	2.110	2.323	17	2.624	2.963
10	2.176	2.41	18	2.644	2.987

height is within the dynamic range of the driver, with the increase of the height, the visual influence on the driver should be increased, but beyond this field of vision, with the height increase, it should tend to be flat. Therefore, the height index and the score should show a certain logarithmic relationship.

Different tree shapes will directly affect the driver's degree of gaze on the crown. Through research, when the tree is spherical, it will give the driver a sense of overall uniformity, and when it is a spire, the radius of the top-down tree is Changes are constantly taking place, so that the driver's degree of gaze will be more spherical. Therefore, it can be inferred that there is a linear relationship between the tree features and the score values.

The crown index of the green belt should be considered in conjunction with the road width. The effect of the width on the crown width is reflected in the fact that when the road width is large, the size of the crown width appears within the driver's field of vision with a small proportion of the road width. At a smaller scale, the crown size is in the driver's field of vision with a large proportion of features.

The proportion of the sky in the green belt is generally between 1 and 50%. The quantification of indicators is not large, and the simple one-way linear relationship can be used directly to describe the relationship between this indicator and the driver's visual impact.

When the visible value of the trunk is larger, it indicates that the tree trunk occupies a larger area of the green belt and attracts the driver's eyes On the contrary, the visual attraction will decrease. It can be assumed that the score value is linear with the score value, and the SPSS output result is used to judge whether the fitting degree and significance are in compliance with the regulations.

In summary, the relationship between the hypothesis and the indicator value of the indicator value of the green belt is:

$$y = a \ln x_1 + bx_2 + cx_3 + dx_4 + ex_5 + g \tag{6}$$

where: a, b, c, d, e, and g are all undetermined constants; x_1 is the height value, that is, the distance from the ground to the top of the tree, the unit is m; x_2 is the tree index value, and the tree is the tree The space structure is divided into a sharp tower shape, a spherical shape, and an umbrella shape. It is determined by the product of

the driver's visual recognition distance and the area of the visible part above the trunk of the driver's perspective. The unit is: x_3 is the ratio of crown width to road width, dimensionless. The tree crown is the average of the north-south and east-west widths of the trees; x_4 is the ratio of the sky, that is, the ratio of the effective sky area to the driver's total area of view at the first viewpoint of the driver, dimensionless; x_5 is the value of the trunk visibility index, determined by the product of the driver's visual distance and the visible trunk area at the driver's perspective, in units of 10^2 m^3.

(3) Using Regress language to solve the pending constant [10]

Using the input plant index variable x_i in the SPSS software, the score is the dependent variable y_i, and the regression language is used to solve the undetermined constant.

For the experimental green belts' height, crown width, trunk visibility, sky ratio, tree form, five kinds of indicator language codes are as follows.

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT score
  /METHOD=ENTER Height Crown Width Ratio Sky Ratio Tree Form Visibility
  /SCATTERPLOT=(*ZRESID,*ZPRED)
  /RESIDUALS DURBIN HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```

After running the program, the input variables will show as in Table 7.

Among them: Table 8 in the R side represents the degree of fit, R square $\geq$ 0.7 shows a good fit. A Durban-Watson value close to 2 indicates that there is no sequence correlation and this regression is not a pseudo regression. Table 9 shows the coefficient of each unstandardized coefficient corresponding to each variable. The significance of $\leq$0.05 indicates that the independent variable has a significant effect on the dependent variable. VIF $\leq$ 5 indicates that there is no collinearity between the respective variables.

Table 7 Enter/remove variables

Model	Input variable	Removed variables	Method
1	Tree visibility, height, crown width ratio, sky ratio, tree's form[b]		Enter

Note [a]Dependent variable: impact value; [b]All the requested variables have been entered

Table 8 Model summary

Model	R	R^2	Adjusted R^2	Standard estimate error	Derby Watson
1	.988[a]	.977	.862	.66824	1.612

Table 9 Model indicator coefficient

Model		Unstandardized coefficient		Normalization coefficient	t	Significant	Collinear statistics	
		B	Standard error	Beta			Tolerance	VIF
1	(Constant)	3.202	4.478		.573	.043		
	Height	1.408	2.301	−.006	−.011	.026	.079	1.870
	Crown Width	1.806	.314	1.289	5.102	.047	.359	2.782
	Tree's shape	.997	7.369	.443	.816	.032	.078	2.930
	Sky Ratio	−.831	4.462	−.155	−.671	.015	.432	2.316
	Tree Visibility	.994	.337	.595	2.018	.050	.263	3.797

Note [a]Dependent variable: score value

According to the above requirements, the results were checked and found to meet the above requirements. In order to establish a multivariate function model with five indicators: greenbelt height, crown width, trunk visibility, sky ratio, and tree shape:

$$y = 1.4 \ln x_1 + 3.1x_2 + 1.8x_3 - 0.83x_4 + x_5 + 3.2 \tag{7}$$

3 Model Validation Based on Eye Tracker Test

3.1 Design of Driving Test Scheme for Eye Movement

The model of the Eye Tracker used in this experiment is TobiiProGlasses2, which can capture and extract the information of the eyes, accurately measure the movement of the eyes and record the eye motion parameters. In the green belt to delimit height, crown, tree, sky ratio and tree trunk visibility of the AOI, and then according to the influence degree model obtained by the SBE method and the proportion analysis of the driver's fixation point recorded by the Eye Tracker in the AOI distribution, the distribution of the driver's fixation point at different points is obtained [11].

3.2 Experimental Preparation and Process

(1) Experimental principles and requirements: In order to ensure the authenticity and validity of the test data and the safety of the driving in the course of the

test, it is ensured that the experimental personnel have no prior knowledge of the purpose. At the same time, on the premise that the study does not affect the normal operation and visual search of drivers, a series of safety measures should be prepared [12].

(2) Experimental process: The object of the experiment is to wear an Eye Tracker to make a good view of a series of preparations, driving the vehicle to the preselected road. The observer sits in the pilot and is responsible for the opening of the experimental software. When the target road reaches the target Road, the software is opened and the driver's visual change is recorded; When driving out of the target Road, check whether the software data record is intact, save and close the eye tracker.

3.3 Eye Movement Analysis in ErgoLAB

ErgoLAB package is used to identify the gaze points of the moving direction scene in the course of the experiment and to generate the view distribution map of the gaze area.

(1) The division of AOI

In order to analyze the driver's attention to each index of green belt in the course of driving, the division of AOI is carried out (Fig. 1). The Division of interest areas should be guided by the following principles: The area of interest of tree stem visibility is divided into the first view of the driver, which clearly distinguishes the tree trunk and the area where the leaves are not obscured (Fig. 1, color blocks 3, 9). The area of interest of the crown is divided into the area above the visibility area of the trunk and the outer diameter of the tree extension (Fig. 1, color block 2, 8). The area of interest in tree height is divided into areas above the crown interest area and above the tree height. The tree shape interest area is divided into tree outline in vitro extending to the road area (Fig. 1 color block 1–7). The sky ratio interest zoning is divided into the sky without vegetation occlusion area above the first visual angle of the driver (Fig. 1, color block 5).

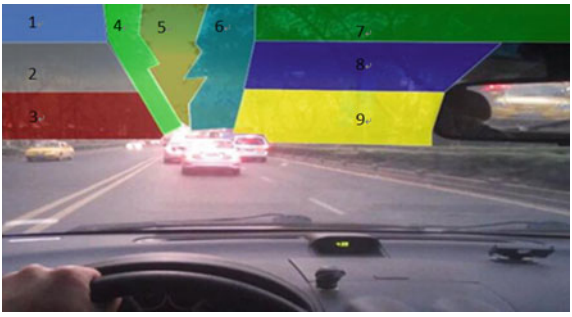


Fig. 1 Division of AOI

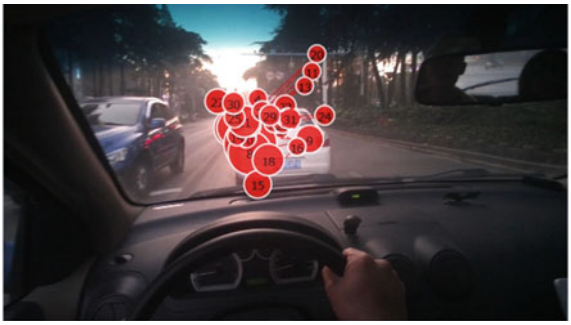


Fig. 2 The track of driver's fixation point

After the division of interest zones, the track map of each section can be obtained through ErgoLAB analysis software. Taking Beijing East Road as an example, the trajectory map is shown in Fig. 2.

(2) Analysis of SBE influence model combined with gaze characteristics

The data were collected by eye tracking system, and the total number of viewpoints and the number of viewpoints in the green belt were obtained based on the number of fixation times. Then the proportion of all fixation points on each road was calculated (Table 10). The average value of all the proportion of all the screenshots was the final result of the road. Finally, the results of each section are fitted with the greening influence values of SBE method. The average value of all the proportions of each screenshot is the final result of the road section. At last, the results of each section are fitted with the greening influence values of SBE method.

According to the statistical results of ErgoLAB software to divide the number of gaze points of AOI, the score of each index obtained by the SBE method and the fitting degree of the experimental data of the eye moving instrument are very high. It is found that the proportion of the fixation point of the driver in the greening belt and the index of the index of the greening belt of the SBE method is on the online relationship (Fig. 3).

The relationship between the fixation point of the driver in the green belt and the proportion of the green belt in the driving process as follows:

Table 10 Road data collation

Road	Fixation point scale	Score values
Beijing East Road	0.2	19.27
Jinxiang River Road	0.3	20.87
Hongwu North Road	0.08	14.06
Taiping North Road	0.05	14.87
Xinmofan Road	0.06	11.71
Changjiang Road	0.5	27.65
Zhongshan North road	0.35	22.39

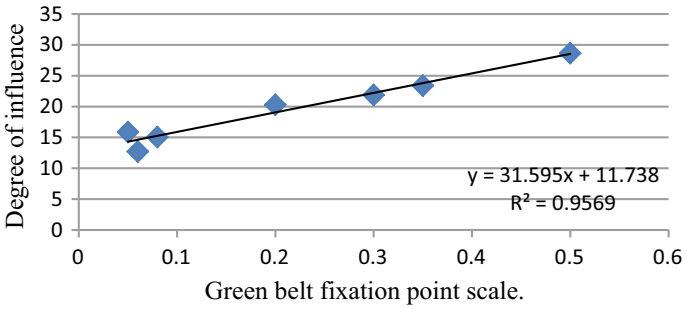


Fig. 3 Proportion of fixation point and score in the linear relationship

$$y = 31.538x + 11.738 \tag{6}$$

where: y is the score of green belt, and x is the gaze of driver on green belt, which accounts for the proportion of total fixation point in driving process.

Based on the objectivity of the scoring system and the analysis of gaze characteristics, the viewpoint ratio of the eye drops recorded in the green belt was verified. There is a linear relationship between the proportion of fixation points of drivers in green belts and the score of SBE method. Therefore, the influence degree of the green belt on the driver's vision can be evaluated according to the value of the green belt model.

3.4 Experiment in Driving Simulator Dome Under Variable Traffic Flow Condition

Under the condition of different traffic flow, the driver travels at the same road at different speeds. As the driver's dynamic vision changes with the change of speed, the visual influence of the roadside green belt on the driver is different [13]. Combining the VR technology with the driving simulation module and using the VR technology to design the scene of different road greening belts, the experimental scheme for the driver to wear eye moving instrument for the gaze collection at different speeds is set up, and the influence of speed on the driver's vision is obtained [14]. According to and the limit of the maximum speed of the road by "Code for design of urban road engineering "and the opinion of many experts, the speed of the experiment is 40, 45, 50, 55 and 60 km/h.

(1) The Study on change characteristics of green belt gaze at different driving speeds

10 drivers were selected to take part in the experiment. The eye movement instrument was used to record the eye movement data of the same road at the speed of two minutes respectively at the speed of 40 km/h, 45 km/h, 50 km/h, 55 km/h and 60 km/h on the driving simulator. At the end of the experiment, the number of gaze points of 10

Table 11 The proportion of fixation point

Experimental speed (km/h)	Proportion of gaze points in green belt
40	0.29
45	0.27
50	0.26
55	0.22
60	0.18

experimenters who dropped on the green belt at different speed for two minutes was analyzed by ErgoLAB software, and the average value was obtained (Table 11).

(2) Setting up a grading model of green belt with different speed

The change of the number of gaze points on the same green belt index was studied. According to Fig. 4, under the same green belt road conditions, the number of gaze points on the green belt of drivers decreases. On the premise of determining the linear relationship between the green belt scoring model and the proportion of the driver's view on the green belt, the relationship between speed and the point of view of each green belt is obtained by the experiment of driving simulation module, and the speed V is taken as a variable into the greening belt scoring model.

It is the same as the above principles. After introducing the speed variable into SPSS, the multivariate function model based on five indexes of road greening belt height, crown width, tree trunk visibility, sky ratio and tree shape can be obtained at different speeds.

$$y = 1.4 \ln x_1 + 3.1x_2 + 1.8x_3 - 0.83x_4 + x_5 - 0.17vs + v_x \tag{8}$$

where: the v_s is designed for the speed of the road. Its value is shown in Table 12.

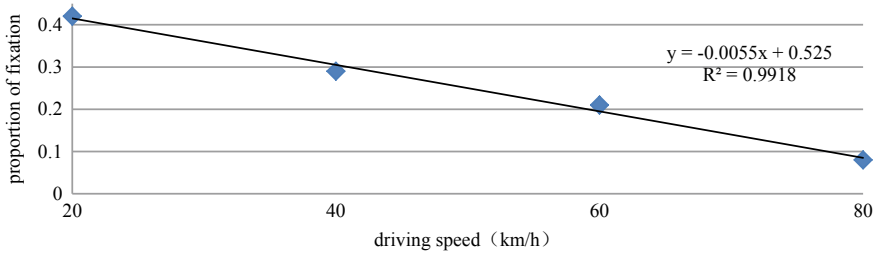


Fig. 4 Relationship of the proportion of fixation point and the speed

Table 12 v_x value proposal table

v_s (km/h)	v_x
40	10.00
45	10.85
50	11.70
55	12.55
60	13.40

4 Grading the Impacts of the Green Belt

Through the above process, the index values of the green belt are substituted into the evaluation model, and the influence value of the green belt is evaluated. However, it is necessary to establish a certain evaluation level for the construction of the green belt. It means that the green belt has different visual effects on the driver in different degree of influence. The evaluation level can be divided into five from the influence to almost no impact. The evaluation level is based on the expert scoring method and the driver's scoring method [15]. In order to make the evaluation standard have a strong universality and accuracy, a large number of road landscape data should be collected, and the drivers of different driving age, gender, height and eyesight should be invited to score. The driver rating method is a score that the driver graded the indicators in the test process. The scoring criteria are 5 points, including 5 points, 4 points, 3 points, 2 points, 1 points [16]. The higher the effect, the higher the score. After the grading, the indicators of all the experimental roads are subdivided. For example, the sub-categories of height can be 0–5, 5–10, 10–15, 15–20 m, etc. The tree form is classified as steeple—type, umbrella—type, ball—type, etc. The different values are grouped into the corresponding sub-categories, and the mean is taken after the generalization. Enter the calculated values in the hierarchy analysis software, the weight of the parent index can be obtained.

The expert grading rule is to invite experts with relevant work and research experience to make appropriate evaluation on the indicators of the green belt selected by the road. The evaluation criterion is made of ten points. From the aspects of visual effects, when the driver thinks the green belt is reasonable, the score ranges from 1 to 3. The indicator has an excessive influence on the driver's vision, and it is recommended to be trimmed with a range of 4–6 points. The indicators have a great influence on the driver's vision, and there are safety hazards and need to be repaired, The score ranges from 7 to 10.

After the above steps, combine the weights of indicators and the expert scoring method to perform metical-operations. The comprehensive evaluation index for the green band index is determined.

$$B = \frac{\sum (X \times F_i)}{n} \tag{9}$$

where: B is the comprehensive evaluation index; X is the weight of each index; F_i is the score of effective expert; n is the total number of experts. Arrange the B of all roads values in numerical order. Calibration of the visual effects of the green belt. The comprehensive evaluation index of the road with better visual effect is lower, which generally accounts for 15–20% of the total number of experimental roads. The general evaluation index of road comprehensive evaluation of visual effect is medium, which generally accounts for 30–40% of the total number of experimental roads. The comprehensive evaluation index of poor visual effects is relatively high, which generally accounts for 15–20% of the total number of experimental roads.

After, enter the road parameter value into the evaluation model. Then, the evaluation interval for different influence degree evaluation is obtained, and these intervals can be used as evaluation criterion to measure whether the green belt construction to be evaluated is reasonable, or whether there is excessive visual stimulus to the driver. General or poor visual effects should be based on the appropriate range of visual effects, select the specific level of each factor reasonably and combine the evaluation model. Make the green belt score fall into the zone, give the driver a comfortable aesthetic feeling and proper alertness and excitement to ensure the stability of the driving.

Through the above experimental procedure., the optimal score range for green belt height, tree form, tree trunk visibility, sky ratio and canopy 5 indicators has been preliminarily obtained from 14 to 19 points. The average score ranges from 11 to 14 and 19 to 23, while the remaining influence values in the above interval are not good visual effects. Therefore, it is necessary to adjust the specific level of each factor reasonably, so that the green belt score falls to 14, to 19 points. But, due to the limitation of time and experimental equipment, and the influence of seasonal on road vegetation, the score interval needs to be further refined and optimized.

5 Conclusion

Through the experiment, it is verified that the visual characteristics of the driver in the process of driving are affected by the road green belt and the traffic flow. Aiming at the characteristics of different factors such as different green belt index and speed control, the visual impact characteristics of the driver were studied, and the green belt model was established with the mathematical analysis. The following conclusions were drawn:

(1) In order to keep the balance between traffic safety and road landscape, the recommended plant height is 15–25 m. The tree, such as sycamore is suitable.
(2) The number of gaze points of AOI in the same road greening area decreases with the increase of vehicle speed, which indicates that different grades of roads should adopt different design standards according to the average driving speed of vehicles and select the specific level of each index, so that the green belt score falls within the range 15–19, this means that the drivers feel comfortable.

References

1. Chen F (2009) The impact of road visual environment. Chongqing Jiaotong University
2. Bai L (2013) Research on construction of city road greening. Appl Mech Mater 2155 (253)
3. Zhou Z, Yuan L, Cui E et al (2012) Analysis of traffic safety by design of urban street median. Highw Eng 37(04):69–72 + 108
4. Shi C, Tian Q, Ge Q, Lu C et al (2014) Application of SPSS statistics software in multiple linear regression calibration data. Contemp Chem Ind 43(06):1112–1113
5. Xiao C, Zhong S, Ma S (2014) Effects of road landscape design on driving behavior. Traffic Inf Saf 32(05):138–145
6. Yu K (1988) Landscape preference: BIB-LCJ procedure and comparison of landscape preference among different groups. J Beijing For Univ 02:1–11
7. Liu W, Gu H, Li C (2011) Expert evaluation method based on Delphi Method. Comput Eng S1:189–191 + 204
8. Yang X, Kang X, Du Z, Bao Y (2012) SBE method-based forest landscape aesthetic quality evaluation of Changbai Mountain. J Northwest Agric For Univ (Nat Sci Ed) 40(06):86–90 + 98
9. Zhang S (1993) The choice of abnormal data: Grubbs Method. Finance Trade Res 03:74–75
10. Chen X, Liang L, Xu G, Liu D (2013) Feature extraction of Kernel regress reconstruction for fault diagnosis based on self-organizing manifold learning. Chin J Mech Eng 26(05):1041–1049
11. Zhang L, Li H, Ge L (2009) Tobii eye tracker applications in human interaction. Chin J Ergon 02:67–69 + 39
12. Lu S (2011) Study on the driver's eye movement fixation characteristics at highway super long tunnel section. Chang'an University
13. Li Q-Y, Zhang Q-P, Wei Q (2015) The influence of driving speed on urban road greening design method. China For Sci Technol 29(03):144–148
14. Zhang J, Jiao X, Chai S (2013) Design and realization of driving simulator system based on virtual reality. Mach Des Manuf 04:39–41
15. Li K (2005) Using analytic hierarchy process in urban road landscape evaluation. J Wuhan Univ (Eng Sci) 01:143–147 + 152
16. Zu B, Guo Y, Huang H (2012) Evaluation of highway landscape sensitivity based on AHP. Environ Sci Technol 35(04):175–178

Study on Urban Road Network Capacity Based on Self-organized Criticality

Zhenlin Wei, Shilong Li, Ailing Huang and Jing Han

Abstract With the rapid growth of the socioeconomic, urban traffic congestion has become acute. In this paper, we studied the internal mechanism of congestion by analyzing the road network capacity, and conducted a qualitative and quantitative research on road network capacity of Beijing. By applying the theory of self-organized criticality, we analyzed the running status data and traffic congestion index of the road network of Beijing from 2008 to 2012, concluding that the road network of Beijing has the power-law characteristics under the relevant scale, following this, the concept of road network capacity based on self-organized criticality has been proposed. Moreover, we established a calculating model of road network capacity based on sandpile model and the theory of cellular automata, simultaneously made a simulation experiment of sandpile model for a specific road network of Beijing. The obtained results show that our model can be applied as an effective approach to calculate the capacity of road network.

Keywords Traffic engineering · Road network capacity · Self-organized criticality · Sand pile model · Cellular automaton · Traffic congestion

Z. Wei · S. Li (✉) · A. Huang · J. Han
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
e-mail: 16114243@bjtu.edu.cn

Z. Wei
e-mail: zhlwei@bjtu.edu.cn

A. Huang
e-mail: alhuang@bjtu.edu.cn

J. Han
e-mail: hanjing19881010@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_38

1 Introduction

The road network is an important urban infrastructure, with which places can be connected and built-up areas can be served [1]. However, With the growth of urban population and vehicles, traffic congestion is becoming more and more serious in many cities, especially in the megacities, like Beijing [2]. Hence inhabitants have to face the environmental pollution, time loss and economic loss caused by traffic congestion. Particularly the road network is a complex system, which is prone to paralysis when the external environment changes, such as a snow, a traffic accident, a traffic control and so on, besides a small disturbance can lead to traffic congestion, even the paralysis of road networks in some cases. Hence it is essential for the stable operation of the road network to study the internal mechanism of traffic congestion.

Road network capacity (RNC) is an important index for measuring the carrying capacity of road network, which also determines the reliability of road network. We studied the RNC by applying the self-organized criticality theory, which will provide a theoretical support for making an intervention to prevent traffic congestion when the road network system reaches the critical state. Therefore, it is of highly significance to study the RNC, which will help to improve the carrying capacity of the road, maintain the reliability of the road network and reduce the probability of traffic congestion [3, 4].

There are a large amount of literatures on RNC, and scholars around the world have done a certain amount of researches on the RNC. They proposed different concepts of the RNC based on different theories and the research will go deeper with time going by. In the early 1960s, Louis et al. putted forward a concept called the consumption of the city's time and space, which considered the RNC by envisioning the city as a container with the properties of time and space [5]. In 1956, Ford et al. proposed the network flow model, and found the algorithm to get the maximum flow of the network, ultimately they put forward the Maximum Flow Minimum Cut Theorem to solve the network [6]. In 1970s, Yamamura et al. published their papers about the largest RNC, in which a method by using traffic assignment to calculate the RNC is put forwarded [7]. Also many scholars have carried on research on RNC in recent years. Kuang et al. supposed that link capacity follows a truncated normal distribution and developed a travel time reliability-based capacity reliability evaluation model, simultaneously a Monte Carlo simulation based solution algorithm is introduced to solve the model [3]. Zhang et al. made a RNC research based on transfer hubs, and traffic modes was added into a road network which is a three dimensional virtual traffic network, finally they found that capacity reliability of road network is inversely proportional to transfer time [4]. Kenneth et al. studied road capacity from the perspective of road investment, and made a model to analyze the optimal balance between capacity and free-flow speed [8]. Chiou et al. researched maximum RNC while taking into account signal-controlled, and proposed a new hybrid search heuristic to exploit the maximum RNC using delay-minimizing signal settings [9]. Christoph et al. addressed the problem of congestion caused by long-term road works, and found that reduced number of lanes, narrower lanes and roadway switching will reduce the capacity of road, eventually

proposed a "Traffic Change" system which has beneficial effects on road capacity [10]. Alexandr et al. applied an approach based on the concept of congestions to calculate traffic capacity of signaled intersections [11]. Tang et al. studied the impact of the road capacity on traffic flow by developing a macro traffic flow model that taking into consideration road capacity, obtaining that the road capacity destroys the stability of uniform flow and produces stop-and-go traffic under a moderate density [12]. Vincenza et al. made a dynamic assignment model to analysis road urban transport network capacity, simultaneously explored some of the properties of network-wide traffic flow relationships in a large-scale complex urban street network [13].

The previous literatures mentioned above illustrate the RNC from different theoretical perspectives, but due to the complexity and variability of the transport system, it is difficult to fit the theory with the actual situation. In this paper, we studied the complexity of road network system, and analyzed the relationships among the internal components, additionally applied self-organized criticality theory in traffic flow of road network system, which has opened up a new perspective for the study of RNC through qualitative and quantitative research method, also the concept of RNC which based on the feature of SOC has been put forward. Moreover through establishing the sandpile model, we figuratively simulate the nonlinear interaction between traffic flows, which has made the evolutional process of the traffic flow closer to the actual situation. The significance of this study is that when the road network system reaches the critical state of the organization, making a destroying timely can prevent the occurrence of large-scale traffic congestion to a certain extent. Hence our study provides a new perspective for the research of RNC.

The rest of paper is organized as follows: In Sect. 2, demonstrate the self-organized criticality of road network. In Sect. 3, design the model for solving RNC based on sand pile model. In Sect. 4, make a simulation experiment of RNC model. Finally, Sect. 5 concludes the paper.

2 Capacity of Road Network Based on Self-organized Criticality

2.1 Self-organized Criticality

The concept of self-organized criticality (SOC) that the system moved towards its critical point without the need to fine-tune the system parameters [14], proposed by Barker et al. is widely used to explain the behaviors of extension dissipative dynamical systems [15]. The way of thinking in self-organized criticality theory is taking into account the whole system's characteristics, which is useful to deal with complex systems.

2.2 The Verification of Urban Road Network's Self-organized Criticality

According to the SOC theory, in order to determine whether a system has the features of SOC, there are two following standards to be applied. One is that the system is widely extended dissipative dynamical (i.e., the system is with openness and out of equilibrium with nonlinear interactions and fluctuation phenomena), and the other one is that the field quantity of system exhibits the space-time power-law distribution. In order to judge the self-organizing criticality of road network, we should judge from the above two characteristics. The process of judgment is as follows.

2.2.1 Extension Dissipative Dynamical System

According to the characteristics of the extended dissipative power system, the process of proof includes the following aspects.

(1) The peculiarity of openness. There has information and energy exchange between the road network system within the boundaries and external systems, such as the increase of traffic volume, the introduction of guidance information, etc. It can conclude that the road network is with openness.
(2) The peculiarity of out of equilibrium. The urban road network system has peculiarity of openness, which makes it possible that the system and the external environment exchange material, energy and information reciprocally. Otherwise, to form "flow", potential difference is necessary. In terms of time, the information of social and economic environment on the input of the system, the construction of roads and other infrastructure, and the amount of motor vehicles are changing constantly. Therefore the structure of material and energy of the whole system is far from static. It makes the traffic system unbalance that the spatial and temporal difference of urban road traffic system, which thus forming "flow".
(3) The peculiarity of nonlinear interactions. Traffic flow in different roads has nonlinear interactions that are based on travelers' subjective choices and decisions.
(4) The peculiarity of fluctuation phenomena. When traffic flow reaches the self-organized critical state and traffic control factors are not enough to damage the system's steady state, the vehicles in different road make the system keep at a steady state through the traffic fluctuation.

Therefore, the road network is a complex system with dissipative structure, and is an extension dissipative dynamical system.

2.2.2 The Space-Time Power-Law Distribution of the Field

(1) The idea of analysis

For the object which follows power-law distribution, according to the fractal theory in space and the 1/f noise in time, there should be the following relationship:

$$N = cr^{-D} \tag{1}$$

where r is the scale of the object; N represents the frequency of r; c is a constant; D is the exponent of power-law distribution.

We also can represent Eq. (1) by log-log transform as follows:

$$\lg N = C - D\lg r \tag{2}$$

where $C = \lg c$.

In terms of Eq. (2), we will study the distribution of traffic flow in Beijing's road network based on following procedure:

① Observe the plot of traffic flow on log-log scale and determine if the shape is linear, which is the characteristic signature of power-law.
② Calibrate Eq. (2) by using the measured data, then we can get a linear function.
③ According to the measured data, we determine the significant level, and get the critical value by checking the correlation coefficient significant test table.
④ Test the significance of correlation coefficient for the function. If the significance of correlation coefficient is larger than the critical value, it is considered that the data follows a power-law distribution, or vice versa.

(2) The results of empirical analysis

Here, we will make an empirical study on the traffic flow's SOC feature. There are two sets of data being used to analyze. One is the traffic states of each road in different times provided by Variable Message Signs (VMS) system of Beijing. The other is the sub-period Traffic Performance Index (TPI) that the value varies from 0 to 10 (the higher the TPI, the heavier the traffic), which is released by Beijing Transportation Research Center. According to these data, we make an analysis on the relation between the number of road sections whose TPI is more than 6 (indicating moderate and severe congested) and their frequencies. The experimental data is collected during the period of from 17:10 to 18:30 p.m. of the year 2008 to 2012 and divided into 5 sets according to weekday, weekend and the year. We present the log-log scale plots of sections' frequencies on weekdays and weekends in Figs. 1 and 2, and roughly find that both the weekday's and weekend's distributions of congested road sections approximately exhibit power-law behaviors as the plots show themselves to be nearly linear. Then we test the significance of correlation coefficient for these 5 sets of data respectively (shown in Table 1). It can be seen that each set of data satisfies the significant level of 0.05.

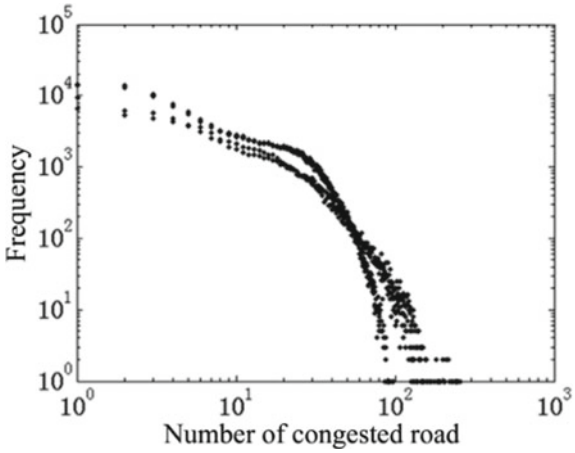


Fig. 1 The relationship between the frequency and number of congested road sections on weekdays from 2008 to 2012

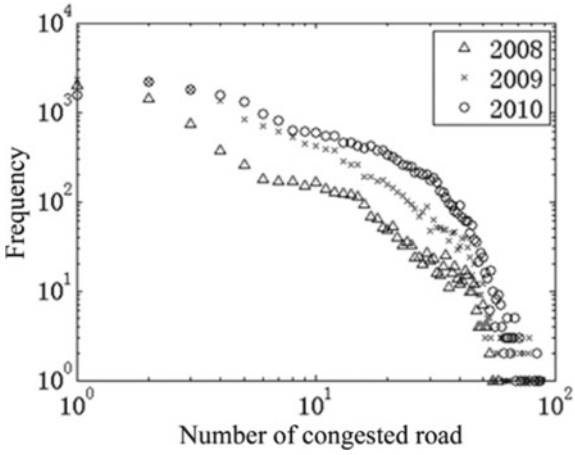


Fig. 2 The relationship between the frequency and number of congested road sections on weekends from 2008 to 2010

As can be seen from Figs. 1 and 2, the number of congested roads and the double logarithmic graph of frequency are close to the line, which follow power-law distribution momentarily. Then the obtained function is tested for the correlation coefficient significance. If the correlation coefficient of the calculation is greater than the critical value obtained according to the actual data, it is considered that the data conforms to the power law distribution characteristics, otherwise no power law distribution characteristics is conformed. Through analyzing those date of urban road network of Beijing, the result is listed in Table 1.

From Table 1, it can be seen that each group of data meets the correlation coefficient significance test whose significance level is 0.05. Therefore, it shows that the number of congested road sections follows the power-law distribution in Beijing's urban road network. Through the analysis, the traffic congestion range and the occurrence frequency conform to the power law distribution when the STATUS is below

Table 1 Test results of correlation coefficient significance

Date	Correlation coefficient	Variance	Distribution function	Significant level $\alpha = 0.05$
Weekdays on 2008	0.7651	36.08	$y = -1.867x + 13.37$	0.754
Weekends on 2008	0.8081	14.71	$y = -1.552x + 9.953$	
Weekdays on 2009	0.7689	33.60	$y = -1.824x + 13.38$	
Weekends on 2009	0.8595	15.34	$y = -1.750x + 10.90$	
Weekdays on 2010	0.7781	28.50	$y = -1.730x + 13.28$	
Weekends on 2010	0.8239	20.22	$y = -1.737x + 11.40$	
Weekdays on 2011	0.8203	22.68	$y = -1.603x + 12.51$	
Weekdays on 2012	0.8863	15.37	$y = -1.606x + 12.33$	

20 and TPI is more than 6. Accordingly the road network is an extension dissipative dynamical system which the traffic flow exhibits the field's space-time power-law behaviors. Hence, we can draw a conclusion that the traffic flow of Beijing's road network can reach a critical state under certain conditions, and then has the SOC.

2.3 The Capacity of Road Network Based on SOC

According to the SOC characteristics of road network, here, we propose the concept of road network capacity, that is: in urban road network systems having dissipative structures, with the increase of traffic flow, the road networks will evolve from the status of normal operation to congestion in a large area, hence, we define the carrying traffic flow by a given size of road network just before it reaches the state of self-organized criticality as the RNC. The capacity of road network based on SOC can be denoted by vehicle kilometer per hour (veh km/h).

3 The Sandpile Model for Road Network

In this section, according to the actual traffic flow in Beijing, we will simulate the congestion process of road network system by the idea of sandpile model and cellular automata (CA), and study the RNC before it reaches SOC state. In the simulation,

Table 2 Similarity comparison between road system and sandpile model

Road network system	CA model based on the concept of sandpile model
Structure and layout of road network	Cellular space
Adjacent roads	Cellular neighbor
Vehicle flow	Grain of sand
Behavior of selecting path for traffic individual	Self-organized behavior of grain
Traffic performance index	Critical height/slope
Road congestion	Sand dumping
Traffic paralysis	Sand collapse

the road network will be mapped into a two dimensional cellular space through the AutoCAD software (AutoCAD is a commercial software package that performs various types of plotting program), then based on CA model, the evolving process of sandpile model will be simulated. Finally, the RNC will be derived from the processed data.

3.1 The Similarity Between Road Network and Sandpile Model

Sandpile mode is the first physical model to display the SOC behavior. This model supposes that the grains of sand are placed onto a plane to form a pile with certain height (slope). With the increase of sand, the pile will reach a critical state that its height (slope) remains unchanged. At this point, the possibility of sandpile collapse triggered by newly added grain of sand increases. The critical state before sandpile collapse is called SOC state.

For constructing road network, CA model is a good way to abstract its form [16]. Meanwhile, as for the evolutional rules of cells in model, the road network system with a large area of paralysis shows a strong similarity to sandpile model. Through analyzing the constitution of road network, as well as CA and dynamic characteristics of sandpile system, we summarize their similarities in Table 2.

3.2 Evolution of Road Network

According to all elements and evolution of sandpile model, we program by JAVA language to implement the model by following procedure (as shown in Fig. 3):

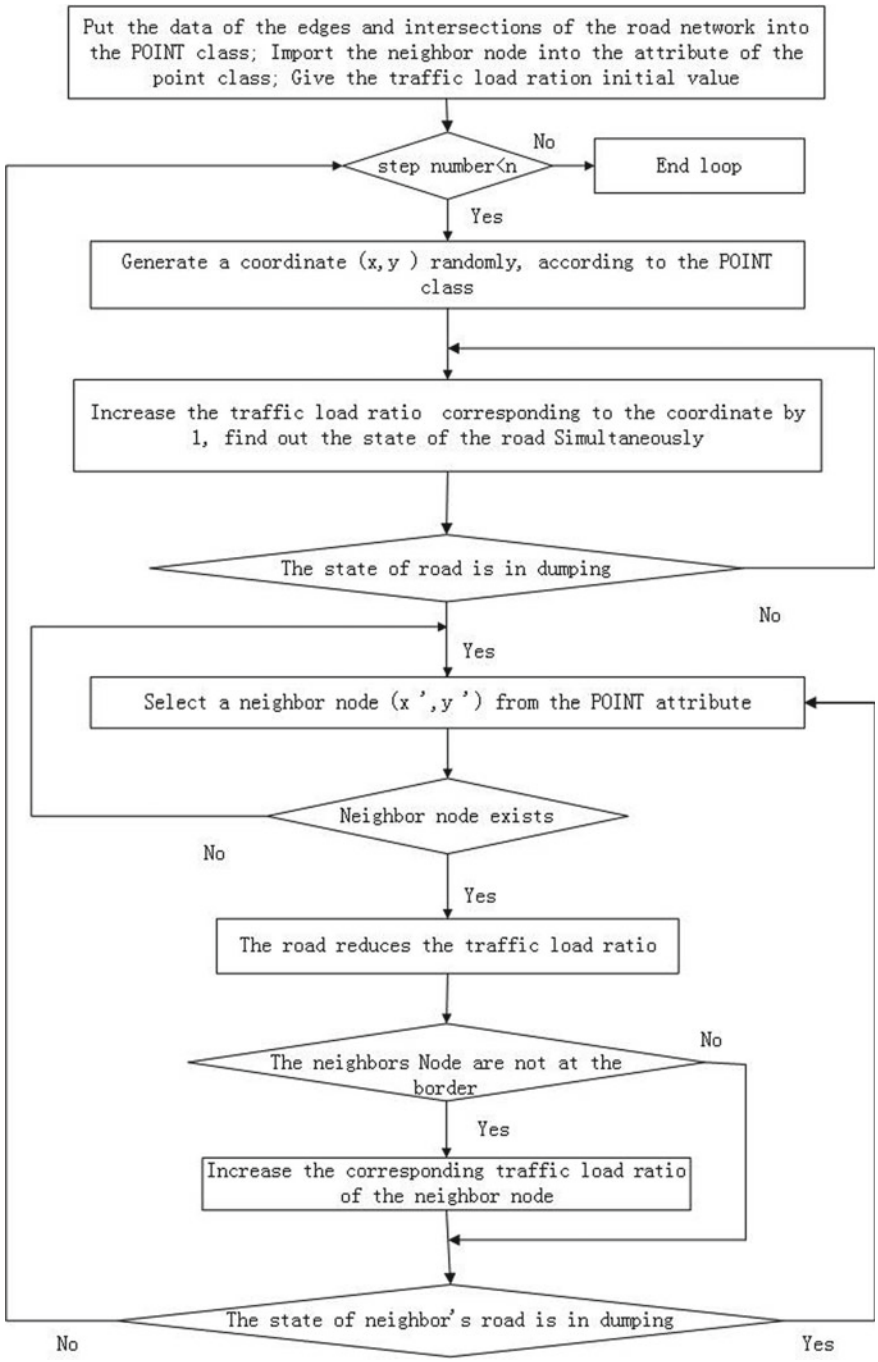


Fig. 3 The procedure of sand pile model to simulate the evolution of road network

(1) Import the cellular space to *POINT* class, and deposit the adjacent edges into the property of $N^2 - N$ edges of *POINT* class, then give an initial value to matrix γ.
(2) Randomly select a road existing in the class of *POINT*, and increase ρ (traffic load ratio) by 0.5 [17], that is, $\rho = \rho + 0.5$; then let $\gamma = \gamma + 1$, γ is twice of traffic load radio.

Traffic load ratio ρ is the ratio of actual traffic intensity to ideal traffic carrying capacity. Generally, it varies from 0 to 3. Here, we assume that the traffic flow of this road remains unchanged and its load ratio keep at 0.5 before next adding of grain. In sandpile model, the increase of road traffic flow can be considered as randomly adding grain to the model.

It should be noted that as each grain of sand is round particle with the same size in the model. However, different roads may have different geometric factors such as grades, width, length, etc., which will influence the traffic capacity of roads with same length. Hence, we need to normalize the traffic flow of each grade road in accordance with the road attributes. This paper uses traffic load ratio to normalize the traffic flow of roads with different grades. For the sake of analysis, we assume that increasing traffic load ratio by 0.5 is equivalent to adding a grain of sand in model.

(3) The average velocity of vehicles on this road can be obtained based on following equation:

$$V = 142.49\mathrm{e}^{-0.4326\gamma} \tag{3}$$

Then the operating state S of the road can be determined according to the average vehicle velocity V:

$$S = 100V/V_d \tag{4}$$

where V_d is the design velocity for the selected road.

According to the value of S, we should determine whether the operating state of a road reaches the congestion status. If not, continue to determine another road. Transform the operating state of road to corresponding probability, and the dumping is carried out. The new traffic flow is transferred to adjacent road according to the corresponding probability, and the traffic flow of original road decreases as well automatically.

(4) The same operation is applied to the adjacent node whose traffic flow changes until new instable nodes disappear.
(5) Repeat the procedure of (2)–(4). If the traffic flow falls outside of the boundary when the collapse occurs in the boundary nodes, it is considered to have escaped the system. Record the number of nodes which are affected by the collapse.

(6) Draw the time series plot of collapse scope, log-log plot of collapse scope and occurrence frequency, time series plot of road network average velocity and relation plot of average traffic load ratio and average velocity.

3.3 The Solving Model for RNC

(1) The definition of critical index TPI

Generally, TPI is measured by the proportion of weighted congested kilometer taking the vehicle kilometer of roads with different grades as weights. Here, the traffic flow is taken into account in our model which is represented by traffic load ratio when calculating the road network's capacity. Thus, we define the critical index as critical congested velocity of road network (Eq. 6), which is calculated by weighting each grade road's congested velocity with taking the length of different grades' roads as weights. According to the empirical analysis of SOC for traffic flow and the TPI level in Sect. 2.2, the maximum speed in the status of moderate congestion for different road grades is regarded as the congested velocity of each road grade. The traffic flow is considered as reaching a self-organized critical state when the velocity reaches the value of critical index.

$$V_c = \sum\nolimits_{i=1}^{n} L_i V_{ci} \Big/ \sum\nolimits_{i=1}^{n} L_i \tag{5}$$

where V_c is critical index, that is, the critical congested velocity of road network; L_i is the length of road grade i; V_{ci} denotes the congested velocity of road grade i.

(2) Critical traffic load ratio

When the traffic flow reaches the SOC state, the critical traffic load ratio of network can be presented based on the critical congested velocity as follows:

$$\rho_c = \sum_{i=1}^{n} \rho_{ci} / n \tag{6}$$

where ρ_c the critical traffic load ratio of road network; ρ_{ci} is the critical traffic load ratio of each road section when the road network reaches the critical state.

(3) The model of RNC

The dominant parts of the ideal RNC include the length, road grade, ideal lane capacity and other factors. The ideal capacity of a lane reaches 1500 pcu/h when vehicle speed is 60 km/h. Besides, due to taking into account intersection reduction factor [18], we distinguish the urban expressway and other roads when calculating the ideal RNC. The ideal RNC is demonstrated as follows:

$$RC = 1500 \times \sum_{i=1}^{n} \lambda_i \mu_i R_i \tag{7}$$

where RC represents the ideal RNC (veh km/h); λ_i is the number of lanes in road i; μ_i denotes the signal intersection (or entrance) reduction factor of road i, in general, 0.41–0.45 for intersection and 0.92–0.96 for entrance and exit of urban expressway; R_i is the length of road i (km).

Thus, the critical RNC CC before its traffic flow reaches the SOC state can be estimated according to the ideal RNC and critical traffic load ratio.

$$CC = \rho_c \cdot RC \tag{8}$$

4 The Simulation of Sandpile Model for Road Network

4.1 The Selection of Network System

This paper selects the road network lying in west of Zhongguancun Avenue, east of Suzhou Street, south of North Fourth Ring, north of Haidian South Road in Beijing as the research object. Within this region, we abstract the expressway, arterial road, minor arterial road and local road as lines with AutoCAD software, and distinguish different road grades with the thickness of line as shown in Fig. 4.

Then we draw a directed network graph (shown in Fig. 5) in which the node represents intersection whose number is the order, and the number of lanes in two directions are labeled on the links (i.e. road sections). There are 62 intersections and 189 directed links in this graph. For intersections on the region's boundary, if there is linked road outside the region, we indicate it with outward arrow.

As this region is one part of Beijing's road network, it can be confirmed as an extension dissipative dynamical system.

4.2 Solution to the Sandpile Model of Road Network

According to CA model, for above selected region, it can be converted as a 62×62 two-dimensional cellular space. Deposit 189 links in the set of adjacent edges. The initial matrix is assigned by value of 0.8185.

Fig. 4 Road network schematic of Zhongguancun West region

4.2.1 Determination of TPI

According to the definition of critical index [2], we can determine the congested velocity of expressway is 35 km/h, arterial road 25 km/h and minor arterial road 15 km/h. Hence, according to Eq. (5), taking the length of roads in Zhongguancun West region as weights, the critical congested velocity of this regional road network is calculated as 17.9278 km/h.

4.2.2 Simulation Result

Carry out the simulation experiment, with critical index that the size of 4000 grains of sand and the average critical velocity of 17.9278 km/h. Finally, we can get the result that diagram of congestion scale sequence (shown in Fig. 6), diagram of double logarithmic plot of the scale and congestion frequency (shown in Fig. 7), diagram of the average speed of network sequence (shown in Fig. 8), diagram of the relationship between the average speed and the average traffic load ratio of road network (shown in Fig. 8) (Fig. 9).

According to Fig. 6, 800 is the critical value that congestion scale has changed dramatically. We also find that the relationship of the double logarithmic plot of the congestion scale and frequency is approximate straight line, next handle the congestion scale according to the power law distribution, and getting the distribution function $y = -2.115x + 8.37$. Then data analysis is conducted, finding that the

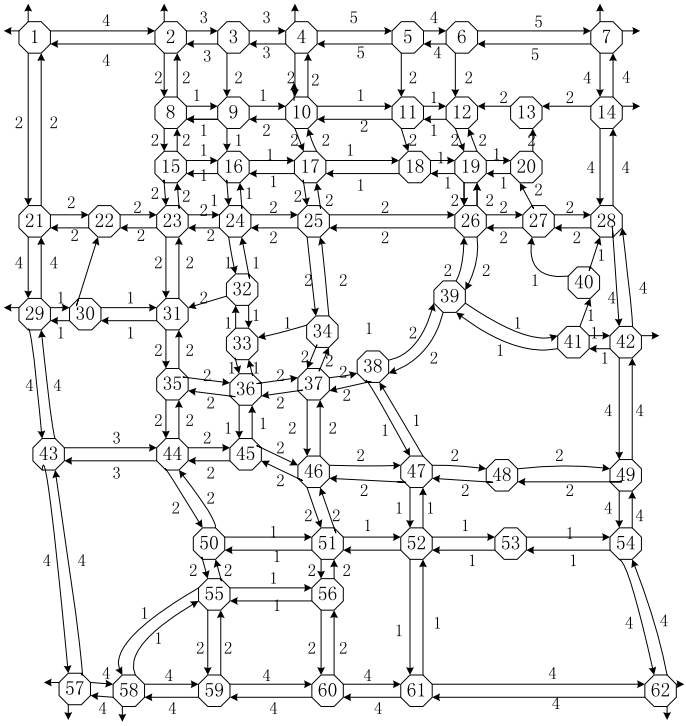


Fig. 5 Two-way road network of Zhongguancun West region

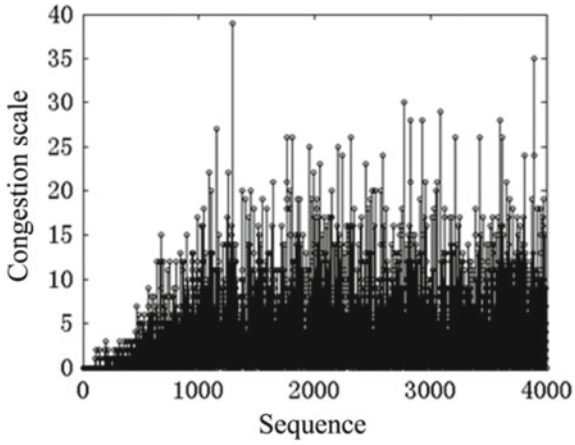


Fig. 6 Congestion scale sequence diagram of 4000

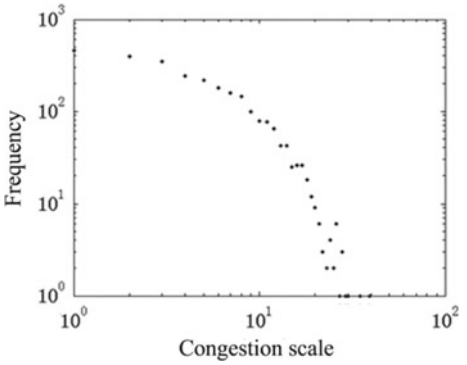


Fig. 7 Double logarithmic plot of the scale and congestion frequency of 4000

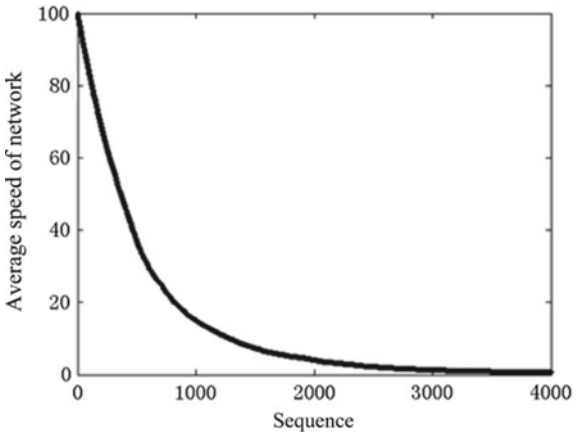


Fig. 8 The average speed of network sequence diagram of 4000

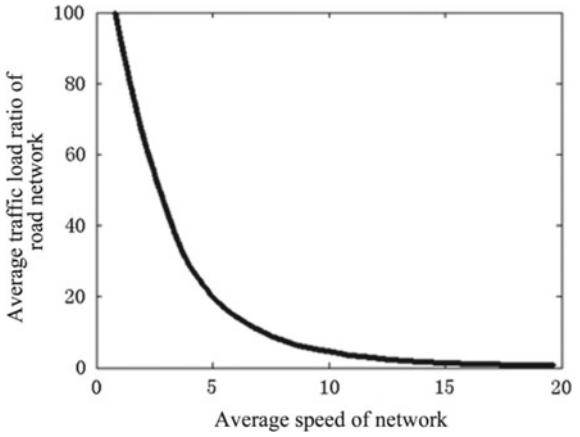


Fig. 9 The relationship between the average speed and the average traffic load ratio of road network of 4000

selected network traffic flow conforms to the power-law distribution, also the system has the characteristics of SOC, when reaching the average critical speed. Ultimately by means of the relation between average speed of road network and average traffic load ratio of road network, the average traffic load ratio of road network is obtained, and then the load capacity of road network is solved.

4.2.3 Determination of Critical Traffic Load Ratio for Road Network

In order to eliminate the influence of experimental scale and times on the results, we carried out 10 experiments for 3 sets of sand size with 4000, 5000, 6000 grains respectively [19], taking the critical congested velocity of 17.9278 km/h as the critical index. Accordingly, the simulated results are shown in Tables 3, 4 and 5 respectively:

Calculate the critical traffic load ratio of road network by averaging all the experiment results, then get $\rho_c = 5.3121$.

Table 3 The experimental results of sandpile model with 4000 grains

Order	Critical congested sequence	Critical traffic load ratio of road network	Variance	Distribution function	Correlation coefficient	Correlation coefficients significant level $\alpha = 0.01$
1	900	5.352934	4.697	$y = -3.736x + 14.16$	0.9312	0.917
2	888	5.326479	2.28	$y = -3.197x + 12.95$	0.9528	
3	891	5.33177	1.11	$y = -2.833x + 12.24$	0.9699	
4	878	5.300024	2.395	$y = -3.129x + 12.85$	0.9486	
5	885	5.294733	4.624	$y = -3.287x + 13.42$	0.9148	
6	901	5.315897	2.572	$y = -3.181x + 12.96$	0.9467	
7	855	5.204786	3.729	$y = -3.358x + 13.4$	0.9323	
8	913	5.395262	1.409	$y = -2.912x + 12.37$	0.9642	
9	886	5.300024	2.314	$y = -3.071x + 12.76$	0.9484	
10	881	5.289442	1.558	$y = -3.182x + 12.92$	0.9667	
Average	887.8	5.311135				

Table 4 The experimental results of sandpile model with 5000 grains

Order	Critical congested sequence	Critical traffic load ratio of road network	Variance	Distribution function	Correlation coefficient	Correlation coefficients significant level $\alpha = 0.01$
1	883	5.310606	2.027	$y = -3.299x + 13.56$	0.9601	0.917
2	877	5.241823	1.437	$y = -2.984x + 12.83$	0.9651	
3	906	5.395262	2.394	$y = -3.209x + 13.35$	0.9509	
4	895	5.300024	3.436	$y = -3.26x + 13.43x$	0.9336	
5	877	5.268278	1.555	$y = -2.988x + 12.84x$	0.9625	
6	879	5.305315	1.171	$y = -3.114x + 13.06x$	0.9736	
7	906	5.421717	1.398	$y = -3.026x + 13x$	0.9677	
8	902	5.416426	3.366	$y = -3.29x + 13.52$	0.9359	
9	899	5.342352	4.263	$y = -3.466x + 13.88$	0.9278	
10	880	5.289442	2.027	$y = -3.097x + 13.13$	0.9551	
Average	890.4	5.329125				

4.2.4 Calculation of RNC

In order to solve the actual RNC [20], besides the critical index and critical traffic load ratio of road network, we also need to calculate the ideal network traffic capacity. Here, the reduction coefficient of intersection is set as 0.43, and that of entrance and exit for related expressway is 0.94. According to Eq. (7), we get the ideal capacity *RC* of road network is 43,121,895 veh km/h. Then the critical capacity *CC* is 229,071,842 veh km/h based on Eq. (8), which means the carrying capacity of the road network before the traffic flow reaches the state of SOC.

5 Conclusion

In this paper, we proposed a methodology to study the RNC based on theory of SOC. Through the qualitative and empirical study, we found that Beijing's road network is

Table 5 The experimental results of sandpile model with 6000 grains

Order	Critical congested sequence	Critical traffic load ratio of road network	Variance	Distribution function	Correlation coefficient	Correlation coefficients significant level $\alpha = 0.01$
1	877	5.26298	2	$y = -3.215x + 13.59$	0.9586	0.917
2	900	5.34764	1.527	$y = -2.933x + 12.96$	0.9619	
3	893	5.32118	2.211	$y = -3.142x + 13.39$	0.9526	
4	876	5.2524	2.45	$y = -3.135x + 13.4$	0.9476	
5	897	5.31589	2.687	$y = -3.363x + 13.89$	0.9499	
6	891	5.36351	0.972	$y = -2.929x + 13.03$	0.9752	
7	877	5.24182	1.375	$y = -3.006 + 13.06$	0.967	
8	882	5.29473	2.04	$y = -3.185 + 13.46$	0.9571	
9	877	5.28944	2.058	$y = -3.033 + 13.2$	0.9526	
10	884	5.27356	2.027	$y = -3.097 + 13.13$	0.9551	
Average	885.4	5.29632				

an extension dissipative system, and the frequency of congested road sections whose TPI is more than 6 exhibits space-time power-law behavior, indicating that with the increase of traffic flow, the road network will reach the state of SOC. According to the SOC characteristics of road network, we proposed the definition of RNC taking into account the traffic operational status. In order to estimate the capacity of road network based on SOC, we presented a sandpile model that integrates the idea of CA model to simulate the evolutional process of road network based on actual data of traffic flow. Additionally we conducted a simulation experiment for a specific road

network of Beijing to verify the validity of the method. The results show that our model can be applied as an effective tool in calculating RNC, which could provide us a theoretic reference to exploit the strategies of controlling and relieving congestion.

Our further research direction is to refine the influencing factors when determining the SOC applicable conditions of road network traffic flow, such as the design speed of road, road grade and traveler's OD. On the other hand, we can also extend our model through adding more factors to make it closer to the reality of the road network.

References

1. Friedricha M (2017) Functional structuring of road networks. Transp Res Procedia 25:568–581
2. He F, Yan X, Liu Y, Ma L (2016) A traffic congestion assessment method for urban road networks based on speed performance index ☆. Procedia Eng 137:425–433
3. Kuang A, Tang Z, Shan L (2013) Road Network Capacity Reliability Considering Travel Time Reliability ☆. Procedia—Soc Behav Sci 96:1818–1827
4. Zhang N, Pan X, Chen F, Fang Q (2013) Road network capacity reliability based on transfer hubs ☆. Procedia—Soc Behav Sci 96:1976–1986
5. Gajjar R, Mohandas D (2016) Critical assessment of road capacities on urban roads—A Mumbai Case-study ☆. Transp Res Procedia 17:685–692
6. Ford LR Jr, Fulkerson DR (1956) Maximal flow through a network. Can J Mathe 8(3):399–404
7. Mcshane W, Roess R. Traffic engineering. Prentice-Hall
8. Small KA, Chen FN (2014) Optimizing road capacity and type. Working Papers 3(2):145–157
9. Chiou SW (2014) Optimal signal-setting for road network with maximum capacity. Inf Sci 273(8):287–303
10. Schwietering C, Feldges M (2016) Improving traffic flow at long-term roadworks ☆. Transp Res Procedia 15:267–282
11. Chubukov A, Kapitanov V, Monina O, Silyanov V, Brannolte U (2017) Calculation of traffic capacity of signaled intersections ☆. Transp Res Procedia 20:125–131
12. Tang TQ, Shi WF, Yang XB, Wang YP, Lu GQ (2013) A macro traffic flow model accounting for road capacity and reliability analysis. Phys A 392(24):6300–6306
13. Torrisi V, Ignaccolo M, Inturri G (2017) Analysis of road urban transport network capacity through a dynamic assignment model: validation of different measurement methods. Transp Res Procedia 27:1026–1033
14. Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: an explanation of the 1/f noise. Phys Rev Lett 59(4):381–384
15. Hwa T, Kardar M (1989) Dissipative transport in open systems: an investigation of self-organized criticality. Phys Rev Lett 62(16):1813
16. Zhao HT, Yang S, Chen XX (2016) Cellular automata model for urban road traffic flow considering pedestrian crossing street. Phys A 462:1301–1313
17. Chen HS, Wu GY (2010) Effects of pair correlation on mean-field theory of BTW sand pile model. Phys A 389(12):2339–2350
18. Afanasyev A, Panfilov D (2017) Estimation of intersections traffic capacity taking into account changed traffic intensity. Transp Res Procedia 20:2–7
19. Cheng R, Ge H, Wang J (2017) An extended macro traffic flow model accounting for multiple optimal velocity functions with different probabilities. Phys Lett A 381
20. Yihu Y (2014) Double-layer ramp-metering model for incident congestion on expressway. J Traffic Transp Eng (English Ed) 1(2):129–137

Transit Signal Priority Optimization for Urban Intersection with the Effects of Downstream Bus Stop

Yuexin Chen, Rui Li, Wei Cao and Changjiang Zheng

Abstract Traffic flow characteristics of intersection will be directly affected by downstream bus service stop, therefore, the effect caused by downstream bus stop should be focused on for optimizing transit signal priority plan of urban intersection. Transit signal priority optimization control unit including signalized intersection and downstream bus stop is identified. Transit signal priority green extension optimization model is proposed, which can optimize transit priority signal timing plan by minimizing total passengers' travelling delay of the whole control unit. Optimized transit priority signal timing plan is determined by using a signal phase allocation method. The proposed model is evaluated using a VISSIM-based simulation platform calibrated with field traffic volume and traffic signal data of Hangzhongmen Boulevard at Beiwei Road and downstream bus stop in Nanjing, China. The results indicate the promising performance (8.38% passengers' travelling delay reduction) of the proposed transit signal priority optimization phasing plans, and the evaluation simulated by VISSIM-based platform also reveal the performance. Consequently, the proposed transit signal priority optimization strategy can significantly improve the traveling efficiency of urban transit system.

Y. Chen (✉)
Institute of Civil Engineering, Nantong Institute of Technology,
Yong Xing Road 14#, Nantong 226002, China
e-mail: 1324982434@qq.com

R. Li · C. Zheng
College of Civil and Transportation Engineering, Hohai University,
Xi Kang Road 1#, Nanjing 210098, China
e-mail: lirui2012@hhu.edu.cn

C. Zheng
e-mail: zheng@hhu.edu.cn

Jiangsu Province Research Center for Traffic and Infrastructure Engineering Technology,
Xi Kang Road 1#, Nanjing 210098, China

W. Cao
Traffic and Transportation Bureau of Haian, Nantong 226600, China
e-mail: 1227343477@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_39

Keywords Public transit · Transit signal priority for intersection · Green extension optimization model · Traffic signal control optimization · Total passengers' travelling delay

1 Introduction

Along with the increasing road traffic congestion of cities, more and more countries begin to actively develop public transport, and improve the traffic priority of public transportation (especially common ground bus) through a variety of ways. Traffic priority signal control, as an important means of bus priority operation, has been applied to road intersections in many urban countries in China. For our country city road intersection, the downstream of many intersections is set up bus stop, this combination will affect the traffic characteristics of intersection of public transport vehicles and social vehicles to a great extent, which will have a certain impact on the intersection of bus priority signal timing scheme. Therefore, we should analyze the influence of vehicle operation characteristics between downstream stops and bus priority signal control intersection, and further study transit signal priority optimization for urban intersection with the effects of downstream bus stop.

Currently, a lot of fruitful achievements have been made in transit signal timing optimization for intersection. Traditional researches view the intersection as the control unit, also, they select some factors as optimization objective, such as the intersection of the vehicle within the vehicle delay at the intersection (social vehicle delay [1], bus delay [2]), passenger delay [3], vehicle arrival punctuality rate [4], the headway [5], and they respectively studied transit signal priority two single handed fork [6–8], arterial coordination control intersection group [9–12], optimization of bus priority signal priority road intersection group [13–15]. In recent years, more and more scholars have begun to study the effect of transit signal timing optimization for intersection caused by bus service stop: Based on the trunk and road network of bus priority signal control environment, Ma [16] and Kim [17] analyze the impact on time station of social vehicles and public transport vehicles caused by bus stops respectively, and optimize transit signal timing optimization which reduce social vehicles delay and transit travel delay for road and intersections; Feng [18] selected the adjacent bus stops between segments (bus stop-to-stop segment) as the control area. Using the method of multiple linear regression, he analyzed respectively the intersection signal delay, traffic flow characteristic parameters, bus stop location and other characteristics which influence the degree of transit operation time and stability of preferential bus, and study bus priority signal timing optimization strategy at intersection under the different traffic scene control section; From the two aspects of traffic efficiency and operation stability, Li [19] did transit signal priority optimization for urban intersection with the effects of bus stop, the effects of bus stops intersection bus priority signal timing scheme is optimized, and compared the impact of bus priority strategy(green extension, red light off early and inserting phase) on the vehicle and bus operation efficiency and stability within bus to bus.

On the basis of previous research, this paper will more focus on the single intersection and downstream bus stop (both as the optimization control area), and the influence of downstream bus stops on upstream priority control intersections is systematically described. From the point of view of optimizing passenger travel efficiency, the total passengers' travelling delay in the control area is taken as the optimal control index of the bus priority signal timing scheme at the intersection. This paper build the green extension optimization model for bus priority at the intersection, the final plan is determined by transit signal priority optimization for urban intersection with the effects of downstream bus stop.

2 The Model of Transit Signal Priority Timing Optimization

Considering passengers' travelling characteristics at the intersection and downstream bus, this paper proposes the calculation model of green extension optimization for bus priority phase at intersection (Intersection and Downstream Bus Stop) to minimize total passengers' travelling delay.

2.1 The Green Extension Optimization Model for Bus Priority at the Intersection

$$\min(d_i + d_s) \tag{1}$$

Taking into account the running characteristics of the bus at the intersection and downstream station, this paper selects the total passengers' travelling delay as the optimal control index, and constructs the green extension optimization model for bus priority at the intersection (illustrated in Eq. 1).

$$\min(d_i + d_s) \tag{1}$$

d_i and d_s illustrate total passengers' travelling delay at one intersection and the downstream bus stop, the duration green extension for bus priority (t_{ge}) will influence total passengers' travelling delay at intersection and the downstream bus stop. In the optimization process, degree of saturation of general vehicle for every entrance at one intersection 0.9 view as the constraint condition of the optimization model.

2.1.1 The Model of Passenger's Delay at the Intersection

There are great differences in the number of passengers carrying buses and social vehicles at intersections, Therefore, according to each phase of signal control intersection vehicle delay, in the calculation of the intersection of passengers' travelling delays, the delay of social vehicles and public transport vehicles should be calculated separately. In the process of studying the traffic characteristics of intersections, it is assumed that there is no lane changing behavior at the entrance of intersection.

With the help of the road traffic manual (Highway Capacity Manual, Transportation Research Board of the National Academies, Washington, D.C., 2000.) method, the average vehicle delay in-Phase of signal control intersection is calculated (illustrated in Eq. 2).

$$\bar{d}_i^j = \frac{0.5C\left(1-\lambda^j\right)^2}{1-\min[1,x^j]\lambda^j} + 900T\left[\left(x^j-1\right)+\sqrt{\left(x^j-1\right)^2+8ex^j/CAP\cdot T}\right] \tag{2}$$

In Eq. (2), $\bar{d}_i^j$ represents average delay of general vehicle for phase j at one intersection; C is cycle length; λ^j, x^j shows green time ratio and degree of saturation for phase j at one intersection; Single intersection signal control type correction coefficient is e; CAP and T symbolizes the lane capacity at one intersection and persistent time.

Green extension time t_{ge} is increased in each phase after optimization, the green time ratio is shown in Eq. (3).

$$\begin{cases} \lambda^p = (g_0^p + t_{ge})/(C-L) & j = p \\ \lambda^j = [g_0^j - t_{ge}x^j/(\sum x^j - x^p)] & j \neq p \end{cases} \tag{3}$$

λ^p, x^p shows green time ratio and degree of saturation for bus priority phase p at one intersection; g_0^j and g_0^p are green time before optimization for bus priority phase j and p at one intersection; L is the lost time.

Combined with the social and vehicle flow and passengers carrying capacity of each entrance road, the passenger travel delay at intersection is calculated, it is as shown in Eq. (4).

$$d_i = \sum_{j=1}^{J} \bar{d}_i^j \left(q_t^j o_t + q_b^j o_b\right) \tag{4}$$

J is total signal phases at intersection; q_t^j and q_b^j represent traffic flow of social vehicles and general vehicles for phase j; o_t and o_b represent the average number of passengers carrying vehicles and buses in the intersection.

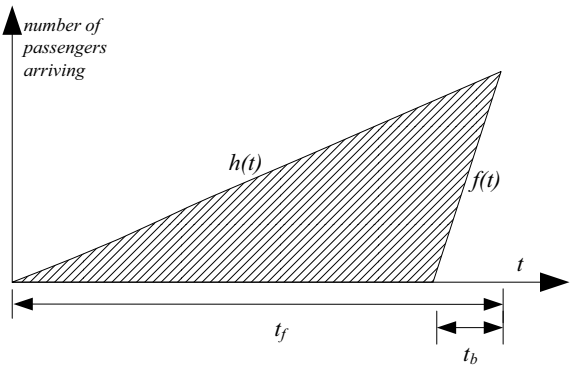


Fig. 1 Visualization of passenger arrival and boarding patterns at bus stop

2.1.2 The Model of Passengers' Travelling Delay at Bus Stop

The passengers' travelling delay (d_s) (illustrated in Eq. 5) in the bus station includes: Passengers' delay waiting for buses (d_s^p) and waiting delay of passengers in car at bus stops (d_s^b).

$$d_s = d_s^p + d_s^b \tag{5}$$

Passengers waiting delay for the arrival of the bus stop in the station can be analyzed by graphical method by analyzing the waiting process of passengers arriving at the station. Figure 1 shows the process of waiting and boarding of passengers at the stops, in which the shadow area is the waiting delay caused by passengers waiting at the stops. $h(t)$ and $f(t)$ is arrival rate of passenger waiting for bus #n of route #m at bus stop and the boarding rate for bus #n of route #m at bus stop.

In order to facilitate the calculation, it is assumed that passengers arrive at the stops at the same time in the unit time, and the bus boarding process also follows the uniform distribution after the arrival of the bus. Passengers waiting delays for bus at the stop at unit time are as shown in Formula (6).

$$d_s^p = \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{h(t_f^{mn})}{2} (t_f^{mn} - t_b^{mn}) \tag{6}$$

M is the number of bus lines and N is the frequency of one bus line at the bus stop. t_f^{mn} shows the following time of bus #n of route #m at stop and t_b^{mn} shows the boarding time of bus #n of route #m at stop.

The passengers' waiting delay during the passengers on board and off process, mainly depends on the dwell time at the bus stop. It is assumed that the bus can be used for passengers to get on and off after arriving at the bus stop (The number of buses arriving at the bus stop is not greater than the parking number of the parking station at the same time, or, the bus does not need to wait in and enter in the station), the dwell time at the bus stop ($t_{d\max}$) for passengers on time (t_b) and off time (t_a) is

the larger value. The passengers' waiting delay in bus after bus stopping in the unit time is shown in Eq. (7).

$$d_b^p = \sum_{m=1}^{M} \sum_{n=1}^{N} \left[\max\left(t_a^{mn}, t_b^{mn}\right) \cdot o_b^{mn} \right] \tag{7}$$

t_a^{mn} is Passenger alighting time of bus #n of route #m, o_b^{mn} is the number of passengers in buses of bus #n of route #m.

2.2 The Model of Traffic Signal Timing Optimization

Using prior optimization to determine the green extension time of bus priority at intersections, the signal timing scheme of the intersection is adjusted under the assumption that the signal period of the intersection is kept constant, which calculates the green phase time of each signal phase (illustrated in Eq. 8). Thus, the model of traffic signal timing optimization at the intersection considering the influence of the downstream stops is determined.

$$\begin{cases} g^p = g_0^p + t_{ge} & j = p \\ g^j = g_0^j - t_{ge} x^j / (\sum x^j - x^p) & j \neq p \end{cases} \tag{8}$$

g^p is green time of phase j after optimization at the intersection; g^j is green time with green extension of phase j.

3 Example Verification

3.1 Traffic Basic Data

This paper selects "Hanzhong Gate Street-North Wei Road" in Nanjing as the intersection after bus priority signal timing optimization and studies the influence of its downstream stop "Hanzhoung Gate Street-Nenjiang Road" for bus priority signal timing optimization scheme, and the optimization scheme of bus priority control signal timing is determined. The layout of the intersection and the downstream station and the signal timing scheme of the intersection are shown in Fig. 2. Hanzhoung Gate Street is the main road for bus priority, bus signal priority at the intersection will be to ensure that straight buses (#9 Road, #82 Road, #552 road bus priority) in the phase 1 pass through the intersection.

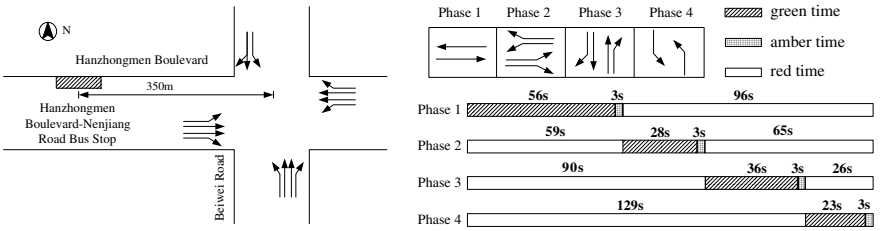


Fig. 2 Layouts and phasing plan for the intersection of Hanzhongmen Blvd and Beiwei Rd

Table 1 Traffic volumes for the intersection of Hanzhongmen Blvd and Beiwei Rd

	Social vehicles/veh			General vehicles/veh		
	LT	TH	RT	LT	TH	RT
WB	193	1008	85	6	16	0
EB	98	796	188	0	20	10
SB	105	182	98	9	0	0
NB	135	208	86	15	9	5

On November 2016, 15–17 days, the vehicle flow of each import Road in evening peak (17:30–18:30) at the intersection is surveyed, the average traffic flow data of three successive days of each import road are as shown in Table 1.

At the same time, the arrival data (illustrated in Table 2) of bus in the stop at night peak in November 15th was selected to study the total passengers' travelling delay at the station. Through field observation, the average number of passengers in vehicles in the region was about 1.8 persons per vehicle, and the average number of passengers in buses was about 28 persons per vehicle.

3.2 Signal Timing Optimization of Bus Priority

With the signal timing priority optimization method at the intersection, the bus priority signal timing scheme is optimized for the intersection of Hanzhoung Avenue and North Wei Road, and the optimized signal timing scheme is shown in Fig. 3. Bus priority green extension time is 13 s.

Comparing the passenger travelling delay between Hanzhong Avenue-North Road intersection and Hanzhong Avenue-Nenjiang Road bus stop before and after optimization (as shown in Table 3), Within the range of the optimized intersection, the passenger travelling delay of social vehicles is reduced by 38,547 s (the delay is reduced by 8.47%), and the delay at the bus stop is reduced by 1273 s after the optimization (the delay is reduced by 6.22%). For the whole optimization control

Table 2 Dwell traffic patterns of transit at Hanzhongmen-Nenjiang bus stop

Bus route	Arrival time	Debarkation/per	Boarding/per	Dwell time/s	Off time
#9	17:37:37	1	3	8	17:37:45
#552	17:46:11	2	3	7	17:46:18
#82	17:46:14	0	5	14	17:46:28
#9	17:46:56	0	2	6	17:47:02
#552	17:57:46	0	1	4	17:58:00
#9	17:57:54	3	2	6	17:58:00
#82	18:02:58	1	3	8	18:03:06
#9	18:09:34	5	5	14	18:09:48
#82	18:09:39	2	2	14	18:09:53
#552	18:16:36	1	5	13	18:16:49
#82	18:21:25	2	4	11	18:21:36
#552	18:21:29	1	2	10	18:21:39
#9	18:23:44	0	3	8	18:23:52
#82	18:25:53	4	3	9	18:26:02
#9	18:28:22	1	2	5	18:28:27
#552	18:30:36	0	1	3	18:30:39

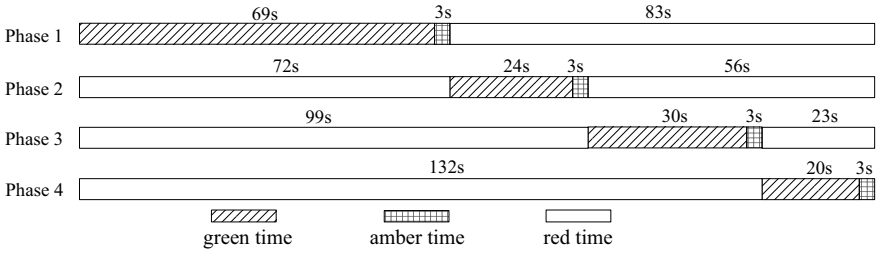


Fig. 3 Optimized TSP plan for the intersection of Hanzhongmen Blvd and Beiwei Rd

Table 3 The comparison of TSP plans for passenger delay reduction ratio

	Passenger delay at intersection/s			Passenger delay at downstream stop/s		
	In social vehicle	In general vehicle	Total delay	Waiting for bus	In general vehicle	Total delay
Present situation	312,010.6	142,921.9	454,932.5	16,557	3920	20,477
Optimized situation	280,540.4	135,845.1	416,385.5	15,358	3846	19,204

Fig. 4 VISSIM-based simulation platform interface

area (intersection and downstream station), the total passengers' travelling delay is reduced by 39,820 s (the delay is reduced by 8.38%).

3.3 Simulation Verification

With the help of VISSIM simulation software, combined with the traffic characteristic parameters of the bus stop of Hanzhong Avenue-North Wei Road intersection and its downstream "Hanzhong Avenue-Nenjiang road" bus stop in Nanjing, the simulation environment (as shown in Fig. 4) is set up.

Through simulation, the delay of the intersection before and after the optimization of the bus signal timing priority scheme is obtained. Combined with the number of passengers carried both social vehicles and general vehicles, based on the VISSIM simulation environment, the total passengers' travelling delay is 520,478.7 and 481,876 s in the whole optimization control area (intersection and downstream station) before and after optimization. Based on the VISSIM simulation analysis, the passengers' travelling delay is reduced by 38,602 s (delay reduction by 7.41%). In spite of the simulation analysis process, compared with the original scheme, the total passengers' travelling delay reduced by the optimized scheme is reduced, however, its optimization effect (7.41%) is enough to show the effectiveness of the optimization scheme in bus signal timing priority optimization at intersections.

4 Conclusion

This paper focuses on the optimization of the bus signal timing priority at the intersection under the influence of the downstream station. Through the characteristics analysis of the intersection and the downstream station, the minimum total passengers' travelling delay is taken as the control objective at the intersection and the downstream station, which optimizes the green extension time of bus priority at the intersection, and then, the bus signal timing priority scheme of the intersection considering the influence of the downstream station is determined. Combined with the actual data of the bus stop of Hanzhong Avenue-Nenjiang road and Hanzhong Avenue-North Wei Road intersection in Jiangsu, Nanjing Province, the bus signal timing priority scheme is analyzed before and after optimization, and build simulation platform based on VISSIM software to simulate and verify the results of the calculation results. The calculation results show that after the transit signal timing priority optimization, intersection and downstream station passengers' travelling delay can be reduced by 8.38%, the simulation shows that the optimization scheme is better in reducing the total passengers' travelling delay.

Acknowledgements This research was supported by the National Natural Science Foundation of China (Grant No. 51508161), Natural Science Foundation of Jiangsu Province (Grant No. BK20140851) and Fundamental Research Funds for the Central Universities of China (Grant No. 2016B08314), and Scientific Research Innovation Program of Graduate Students in Universities of Jiangsu (Grant No. SJZZ16_0086).

References

1. Li M, Yin Y, Zhang W, et al (2011) Modeling and implementation of adaptive transit signal priority on actuated control systems. Comput-Aided Civil Infrastruct Eng 26(4), 270–284
2. Ma W, Liu Y, Yang X (2013) A dynamic programming approach for optimal signal priority control upon multiple high-frequency bus requests. J Intell Transp Syst Technol Plann Oper 17(4):282–293
3. Christofa E, Papamichail I, Skabardonis A (2013) Person-based traffic responsive signal control optimization. IEEE Trans Intell Transp Syst 14(3):1278–1289
4. Furth P, Muller T (2000) Conditional bus priority at signalized intersections: better service quality with less traffic disruption. Transp Res Rec J Transp Res Board 1731:23–30
5. Hounsell N, Shrestha B (2012) A new approach for co-operative bus priority at traffic signals. IEEE Trans Intell Transp Syst 13(1):6–14
6. Dion F, Hellinga BA (2002)Rrule-based real-time traffic responsive signal control system with transit priority: application to an isolated intersection. Transp Res Part B: Methodol 36(4):325–343
7. Christofa E, Skabardonis A (2011) Traffic signal optimization with application of transit signal priority to an isolated intersection. Transp Res Rec J Transp Res Board 2259:192–201
8. Wang Z, Xia L, Luo D (2010) Adaptive transit priority control at isolated intersection. China J Highw Transp 23(4):84–90
9. Dion F, Rakha H, Zhang Y (2004) Evaluation of potential transit signal priority benefits along a fixed-time signalized arterial. J Transp Eng 130(3):294–303

10. Vasydevan M (2005) Robust optimization model for bus priority under arterial progression. University of Maryland, College Park
11. Wang D, Zhu H, Biw Y et al (2011) Bus signal priority method at arterial signal progression. J SE Univ (Nat Sci Ed) 41(4):859–865
12. Li Z, Zhu M, Wang Ba, et al (2015) A transit priority model for arterial roads based on the delays of upstream and downstream intersections. J Transp Inf Saf 33(5):36–42
13. Mesbah M, Sarvi M, Currie G (2009) New methodology for optimizing transit priority at the network level. Transp Res Rec 2089(2009):93–100
14. Mesbah M, Sarvi M, Currie G (2011) Optimization of transit priority in the transportation network using a genetic algorithm. IEEE Trans Intell Transp Syst 12(3):908–919
15. Ghanim M, Abu-lebdeh G (2015) Real-time dynamic transit signal priority optimization for coordinated traffic networks using genetic algorithms and artificial neural networks. J Intell Transp Syst Technol Plann Oper 19(4):327–338
16. Ma W, Ni W, Head L, Zhao J (2013) Effective coordinated optimization model for transit priority control under arterial progression. Transp Res Rec 2356(2013):71–83
17. Kim W, Rilett LR (2005) Improved transit signal priority system for networks with nearside bus stops. Transp Res Rec J Transp Res Board 1925:205–214
18. Feng W (2014) Analyses of bus travel time reliability and transit signal priority at the stop-to-stop segment level. Portland State University, Portland
19. Li R, Jin PJ, Ran B (2016) Bi-objective optimization and evaluation for transit signal priority strategies at bus stop-to-stop segment. Math Probl Eng, Article ID 1054570, 1–12

Speed Limit Analysis for Street in Residential Block Based on Minimum Network Costs

Xiaoning Wang and Shaosha Fan

Abstract With the implementation of government's new guideline on block system, the gated communities will be gradually opened. However, the entrance of a large number of social vehicles into the residential district produces traffic problems such as environmental problems and noise problems. This study proposed a before-after analysis of speed limit for streets in residential block. Traveler time value model, automobile exhaust treatment cost model and traffic noise prevention cost model are considered to minimize the general network costs, and the best speed limit is obtained. Taking the Liaohe district in Harbin as an example, 15 is the initial value and 5 is the interval value, then the five values are 15, 20, 25, 30, 35. The daily total costs of the study area are 298.99, 309.42, 324.46, 326.27, 350.87 yuan correspondingly. 15 is the corresponding value of the minimum total cost. Therefore, it can be set as the speed limit.

Keywords Traffic engineering · Speed limit · District of block system · The cost of network system

1 Introduction

In 2016, China released a central government guideline on urban planning which recommends that the roads in the community established in the future should allow vehicles outside the community to pass, and the existing closed communities should also allow vehicles outside the community to pass. The objective is to ban gated communities and promote the block system in order to make the internal road publicized.

The implementation of the block system will lead to a series of traffic environment and safety issues. The speed limit can not only reduce the probability of accident, but also affect the noise of the vehicle that affects people's good life and the harmful

X. Wang (✉) · S. Fan
Harbin Institute of Technology, Institution of Transportation Science and Engineering, Heilongjiang, Harbin 150090, China
e-mail: 885381345@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_41

gases. Therefore, it is important to study the speed limit for residential streets of the block system. There have been many studies on the road speed limit. Do Duy Dinh built a model to estimate how much the speed of the vehicle on this road would be if the vehicle on the city road could not drive at a speed of more than 30 km/h [1].Charles Goldenbeld analyzed what kind of situation will occur in different country roads which set the same speed limit of 80 km/h, and analyzed the role played by roads, environment and personality traits in forming the above situation [2]. Aarts uses an algorithm to study the different speed characteristics and which speed characteristics can be set to ensure that vehicles on the road travel safely [3]. Cheng Guozhu proposed a method of setting the maximum speed limit of the vehicle on freeway considering comprehensively the operation efficiency, safety, economy and comfort [4].

In summary, the previous research was to achieve the goal of allowing vehicles on the city roads and vehicles on the highway to travel fast and reach safely. However, after the implementation of the block system, residents may pay more attention to the traffic noise the, traffic collisions and the environmental problems in the residential district. Therefore, this study investigates the speed limit model for the residential streets of the block system by considering automobile exhaust and traffic noise. First, this study quantifies the time cost of traveler, the treatment cost of automobile exhaust gas and the setting noise barrier cost. Then the speed limit model is built by taking the minimum cost of the system as the goal.

2 Model Development

The model constructed in this paper takes into account three factors: the time value of travelers, the cost of automobile exhaust treatment, and the cost of traffic noise prevention. The time value of travelers is the time-effectiveness of the traveler; the automobile exhaust cost is the cost required to control motor vehicle exhaust; the traffic noise cost is the cost required to set up noise barriers. This chapter will explain how to calculate the cost of each part.

2.1 Traveler Time Value Model

The total time that spent by road network system travelers is calculated in two parts: road section and intersection which is classified into unsignalized intersection and signalized intersection [5–7]. This paper defines the traveler's time value model as total time times GDP per capita.

$$W = \left[\sum_{l=1}^{n} \left(\alpha T_{l0} \left(\frac{Q_l}{C_l} \right)^{\beta} \times Q_l \right) + \frac{\sum_{d=1}^{m} (Q_d \times y_d) + \sum_{u=1}^{z} (Q_u \times d_u)}{3600} \right] \times C_T \tag{1}$$

where W represents the time value of all travelers in the system (yuan); T_{l0} represents the free travel time; Q_l represents the traffic volume of path l before the block system implemented (veh/h); C_l represents the actual capacity; Q_d represents the traffic flow at signalized intersections before the block system implemented (veh/h); y_d is the average delay of vehicles at signalized intersections before the block system implemented (s); d_u is the average delay at the unsignalized intersections before the block system implemented (s); Q_u represents the traffic flow at unsignalized intersections before the block system implemented (veh/h); The travel time of section is obtained by BPR function and the delay time of intersection is obtained by modified Webster model.

2.2 Automobile Exhaust Treatment Cost Model

The total cost of automobile exhaust treatment is obtained by calculating the quantity of various pollutants discharged by automobile and the treatment cost per unit quantity of pollutants.

$$Q_j = E_j A_a \tag{2}$$

where Q_j represents the amount of pollutant j emitted by the vehicle on road a per unit length (g/km); E_j represents the emission of pollutant j per vehicle per unit length (g/km veh); A_a represents the traffic volume of road a (veh/h).

The cost of automobile exhaust treatment C_v can be obtained by the following Formula (3)

$$C_v = \left(\sum_{j=1}^{3} C_j \cdot Q_j \right) \cdot l_a \tag{3}$$

where C_j represents the treatment cost per unit quantity of pollutant (yuan/g); l_a the length of road a (km).

This paper only considers the exhaust pollution of gasoline engines, and selects carbon monoxide, hydrocarbons and nitrogen oxides as the main pollutants. The selection of E_j value needs to correct adapted to local conditions. For example, the E_j used in the case of this article is based on relevant studies in China and Harbin [8–12]. These results are compared and analyzed to remove significantly larger and smaller values, as summarized in Table 1.

Table 1 The single-vehicle pollutant emission factors applied in this paper

车速	污染物	小型车
15	CO	77.38–83.38
	HC	9.49–11.89
	NO_X	3.13–5.13
20	CO	64.78–70.78
	HC	7.91–10.31
	NO_X	2.74–4.74
25	CO	55.18–61.18
	HC	7.08–9.48
	NO_X	2.53–4.53
30	CO	45.58–51.58
	HC	6.25–8.65
	NO_X	2.32–4.32
35	CO	38.38–44.38
	HC	5.94–8.34
	NO_X	2.22–4.22

2.3 Traffic Noise Prevention COST Model

The traffic noise model of FHWA is selected to determine the range that excess noise influenced. Through the investigation of the residential situation in the range that excess noise influenced, the number of the households needing the noise insulation wall is determined. Then the result is obtained by multiplying the cost of the unit noise insulation window [13, 14].

The FHWA model is shown as follows:

$$Leq_a(d_a) = 10 \cdot \log_{10}\left(0.029 \cdot \frac{\varphi'}{180} \cdot Q_a \cdot J_a \cdot \left(\frac{50}{d_a}\right)^{1+\alpha}\right), \forall a \tag{4}$$

Among them, φ' is calculated by the Formula (5–7), J_a can be calculated by the Formula (8):

$$\varphi' = 2 \cdot \varphi \cdot \left[1 - \frac{M}{|\varphi|} \cdot \left(\frac{|\varphi|}{90}\right)^N\right] \tag{5}$$

$$M = 90 \cdot \left(\frac{0.58 \cdot \alpha^{0.9}}{0.58 \cdot \alpha^{0.9} + 1}\right) \tag{6}$$

$$N = \frac{1}{0.134 \cdot \alpha + 0.225} \tag{7}$$

where φ' is the noise reduction amount caused by incomplete blocking; Q_a represents the volume of path a (vel/h); α is a unitless ground-cover coefficient, as most of the residential area of the ground is very hard, so α is taken as 0.375; J_a is the total noise amount generated by link a; φ is the angle from the noise receptor to the two lines of the road.

$$J_a = \frac{F_a}{1.609 \cdot v_a} \cdot \left((1.609 \cdot v_a)^{4.174} \cdot 10^{0.125} + 10^{G_a}\right), \forall a \tag{8}$$

where F_a is the proportion of different models in the link a, this paper only considers light vehicles, so F_a is taken as 1; v_a is the running speed of the vehicle on path a (km/h); G_a represents the proportion of vehicles indifferent driving states on the road a. Assume that all vehicles on the road of block system are considered light vehicles, so G_a is taken as 5.60.

2.4 The Model of Calculating Road Speed Limits for the District of Block System

This article takes the road network system's time cost, automobile exhaust cost and traffic noise cost weight equally, defines the sum of the three as the total cost [15–18], the resulting model is as follows:

$$Z = \sum_{i=1}^{t} (W_i) + C_v + \frac{C_p}{365} \tag{9}$$

where Z is the average daily travel and environmental costs of road network (yuan); t represents the hours that district of block system opened per day; C_p represents the annual cost of ventilation soundproof windows (yuan).

3 Case Analysis

3.1 Testing site

The location of Liaohe community is shown in Fig. 1. All field data, including traffic speed, traffic volume and traffic composition, were collected from 7:00 AM to 6:00 PM using video camera and radar tachometer.

When calculating the cost of each part, it is necessary to know the road traffic flow, however the traffic flow of the opened community road is unknown. For the sake of research, it is assumed that the implementation of the block system only affects the road network adjacent to the community, that is, the traffic volume that entrancing and

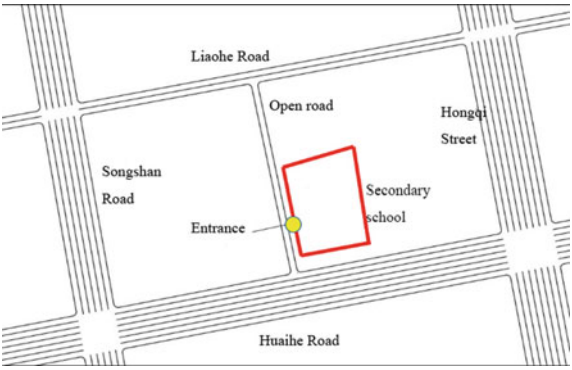


Fig. 1 The Liaohe district

exiting of the community intersection from the external region remains unchanged. According to the current road traffic, the amount of OD between the peripheral nodes of the region is inferred and distributed to the road network where the community road is opened, and the road traffic flow can be predicted. According to street attributes and using the predicted traffic flow to calculate the cost of each part.

The study took five speed limits of 15, 20, 25, 30, 35.

The four intersections around the Liaohe community are all signalized intersections. The intersection between the road within the community and Liaohe Road, the intersection between the road within the community and Huanghe Road are all unsignalized intersections. According to the predicted road traffic volume after the open road, taking Harbin's per capita hourly GDP (6.46 yuan/hour) as the average value generated by the traveler's unit hour to calculate the time cost of the road network.

As can be seen from Fig. 2, under the five speed limit values of 15, 20, 25, 30, 35. When the speed limit is 25 km/h, the time cost is the largest. It is worth noting that although in good traffic conditions, the higher the speed limit is set, the faster the vehicle can travel, the calculations presented here show that time costs do not decrease monotonically with speed. This is because travellers will make corresponding choices in route selection due to different speed limits, thus changing traffic flow, and traffic conditions and travel costs on the road will change.

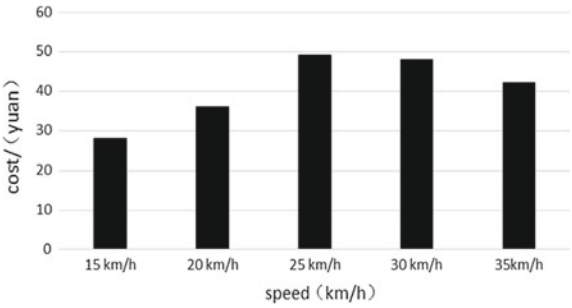


Fig. 2 The travel time cost difference

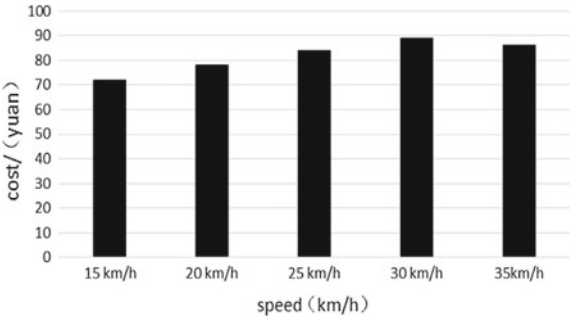


Fig. 3 The traffic accident cost difference

Based on the traffic flow in the community road redistributed after the implementation of the block system, combined with the parameters of Table 1, calculate the total amount of pollutants using Formula (2). The NO_X treatment cost is 8.36 yuan/kg, the CO treatment cost is 1.05 yuan/kg, and the HC treatment cost is 3.34 yuan/kg [19, 20]. The cost of the pollutants discharged from the vehicles in the community road is calculated by the unit exhaust treatment cost multiplied the quantity of the exhaust gas generated by the vehicles running on the road in the community. Figure 3 shows the exhaust cost corresponding to the five speed limit values.

This study solves the traffic noise problem by installing ventilation sound insulation windows for houses on both sides of the street. The ventilation and soundproof windows have low cost and obvious noise reduction effect, which can drop more than 20 decibels. There are 48 households in Liaohe Community who need to install ventilation and sound insulation windows. As we can be seen from Fig. 4, although the range that excess traffic noise effected is different, the number of residential units in the range are the same.

Figure 5 indicates the daily total costs of the study area are 298.99, 309.42, 324.46, 326.27, 350.87 yuan corresponding to five speed limit values. 15 is the value corresponding to the minimum total cost. Therefore, it can be set as the speed limit.

In this case, the open road linear conditions are good. However, in practical applications, safety checks should be performed on the roads with poor line conditions based on line-of-sight. Pedestrian safety is an important factor in the design of roads within a community, and the speed limit policy should be combined with the design of

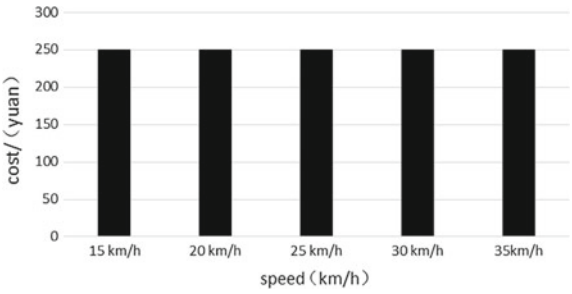


Fig. 4 The automobile exhaust cost difference

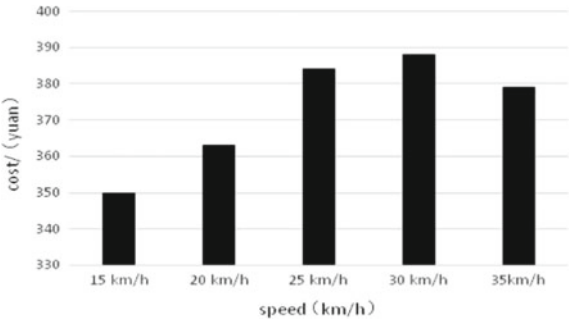


Fig. 5 The traffic noise cost difference

transport facilities. Using auxiliary means such as deceleration signs and deceleration belts at crosswalks and key passages achieves a balance between safety, quietness and efficiency. For example, in this example, as shown in Fig. 1, crosswalks, deceleration signs, and deceleration zones should be provided at the entrance to the secondary school.

In addition, the speed limit is only a small part of the residential district block. Before opening roads within a community, consideration should be given to the linear conditions of the road within the community, and the density of people. Only roads that meet certain requirements can be considered to open (e.g., roads with insufficient width are not suitable for opening).

In this paper, traveler time value model, automobile exhaust treatment cost model and traffic noise prevention cost model are calculated. And then built network cost model considering these three. Taking Liaohe District of Harbin City as an example, it is concluded that 15 is the value corresponding the minimum total cost. Therefore, it can be set as the speed limit. The proposed method fully considers the need for tranquility in residential areas and has a high application value in the implementation phase of the block system policy.

References

1. Dinh DD, Kubota H (2013) Profile-speed data-based models to estimate operating speeds for urban residential streets with a 30 km/h speed limit. IATSS Res 36(2):115–122
2. Goldenbeld C, van Schagen I (2007) The credibility of speed limits on 80 km/h rural roads: the effects of road and person (ality) characteristics. Accid Anal Prev 39(6):1121–1130
3. Aarts L, van Nes N, Wegman F et al (2009) Safe speeds and credible speed limits (sacred speed): a new vision for decision making on speed management. In: Compendium of papers of the 88th TRB annual meeting, pp 11–15
4. Cheng G, Pei Y, Chi L (2009) Method of setting maximum speed limit based on minimum generalized running cost of vehicle. J Jilin Univ 39(4):901–905
5. Yan Z, Du Y, Zhang R (2009) Transport economics. China Communications Press, Beijing, p 93
6. Wang W, Chen X (2007) Traffic planning. China Communications Press, Beijing, p 84
7. Hailong GAO (1998) Delay calculation method of PCE at signalless intersections. J Southeast Univ (Natural Science Edition) 28(3):9–13

8. Fan S, Tian L, Zhang D et al (2015) Study on vehicle exhaust emission factors in Beijing. Environ Sci 7:2374–2380
9. Yu K (2016) Research on vehicle exhaust emission factors and emission test methods. Master Thesis of Zhejiang University of Technology
10. Liang T (2016) Study on characteristics of typical pollutants emissions from vehicle exhaust in Harbin City. Master Thesis of Harbin Institute of Technology
11. Liao Yubo Yu, Zhi Z et al (2012) Study on the characteristics of vehicle exhaust emissions in Guangzhou city. Environ Sci Technol 35(1):134–138
12. Wang X, Zhang H, Fu Y (2016) Research on computing method for amount of emission of motor vehicle exhaust. Technol Highw Transp 32(1):130–134
13. Wang Y, Szeto WY (2017) Excessive noise paradoxes in urban transportation networks. Transportmetrica A: Transp Sci 13(3):195–221
14. Jung FW, Blaney CT (1988) Highway traffic noise prediction for microcomputers: modeling of Ontario simplified method. Transp Res Rec 1176(1):41–51
15. Yang Z, Liu P, Zhang Y et al (2016) Multi-objective analysis of using U-turns as alternatives to direct left turns at two-way stop-controlled intersections. J Adv Transp 50(4):389–405
16. Jiang Y, Szeto WY (2015) Time-dependent transportation network design that considers health cost. Transportmetrica A: Transp Sci 11(1):74–101
17. Long J, Szeto WY, Huang HJ (2014) A bi-objective turning restriction design problem in urban road networks. Eur J Oper Res 237(2):426–439
18. Li ZC, Lam WHK, Wong SC et al (2012) Environmentally sustainable toll design for congested road networks with uncertain demand. Int J Sustain Transp 6(3):127–155
19. Yan Y, Zhang W, Li X et al (2001) Theoretical base and drafting principles of Zhengzhou total pollution charges site scheme. Chongqing Environ Sci 23(3):5–8
20. Tan Y, Zheng S (2007) Research on the determination of environmental pollution unit governance costs. Prod Res 24:52–53

Research on Evacuation Model of Evacuees with Luggage

Shuang Chen, Nan Zou and Fan Wu

Abstract Emergency evacuation in large traffic facilities has been important issues of social security. Current evacuation models less consider the situation, which many passengers carry luggage in large transportation hub. This paper proposes a new cellular automation model according to characteristics of luggage. The distance parameter, spatial parameter, empty parameter, obstruction parameter, luggage following rules and new route created rules are introduced in the model. This paper mainly analyzed the effects of carrying or leaving luggage on evacuated pedestrian flow. Evacuation time is verified based on the simulation under the different space occupancy and different ratio of pedestrian with luggage. The article takes a waiting area in railway station as an example to simulate. The result shows that if moving with luggage, as the initial space occupancy grows, the increase of the proportion of people with luggage has more obvious impact on increase of evacuation time; Besides, when pedestrian space occupancy surpass 60% and the ratio of pedestrian with luggage exceed 50%, due to abandoned luggage loss of mobility, the effect of hindering the subsequent pedestrian movement become stronger. Carrying luggage rather than leaving luggage may make overall evacuation time shorter.

Keywords Systems engineering · Evacuation of pedestrian flow · Cellular automata · Evacuees with luggage · Computer simulation

1 Introduction

In the emergency evacuation process, though most people know that carrying luggage will slow down the movement [1], many of them still insist on carrying luggage.

S. Chen · N. Zou (✉) · F. Wu
School of Control Science and Engineering, Shandong University, Jinan 250061, China
e-mail: nanzou@sdu.edu.cn

S. Chen
e-mail: chenshuangcelia@163.com

F. Wu
e-mail: wf_sdu@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_42

A questionnaire survey of train passengers showed [2] that when an emergency occurred, 69.16% of people chose to carry baggage, of which 11.06% chose to carry expensive heavy luggage, even 5.03% chose to carry all the luggage. There always are large number of persons with carrying luggage in important transportation hubs such as train stations, bus stations, and subway stations. Once an emergency occurs, it can easily cause heavy casualties. Hence, the evacuation of people carrying luggage is worth researching.

In recent years, crowd evacuation behavior has attracted widespread attention in the academic community. Typical research models include social force model, cellular automata model and lattice gas model. The cellular automata model in time and space has a micro-discreteness feature. It can transform the continuous complex process of pedestrian movement into discrete interaction, and the evacuation process of pedestrians can be better simulated by setting appropriate evolution rules.

Zhang et al. [3] introduced absorption coefficient and anti-dead-damage coefficient in the social force model to describe the process of crowded force transmission; Feng et al. [4] introduced attracting parameters, occupying parameters and benefit parameters to simulate the habit of Chinese pedestrian that walking along the right side; Varas et al. [5] solved the pedestrian movement simulation problems when there are obstacles; Fu [6] and Ren et al. [7] found that walking together can lead to congestion by simulating the group behavior; Cheng [8] propose a computational model to research cooperation and evacuation efficiency, the model provided a decision-making basis for emergency response involving large-scale crowd evacuation. Luo [9] compared the four update schemes, namely parallel scheme, random scheme, order-sequential scheme and shuffled scheme and results show that parallel leads to the longest evacuation time and the lowest utilization of exit. In addition, update schemes have the same adaptability to obstacles and complex environments. Fu et al. [10] proposed a multi-velocities floor field cellular automaton model to describe discrepancies in pedestrian movements such as fatigue properties; You et al. [11] and Hun et al. [12] extended the cellular automata model to a multidimensional space, and queue time and critical time are considered in the evacuation strategy. Vermuyten et al. [14] conclude that most recent models clearly include pedestrian dynamics, but more attention should be paid to the calibration and implementation of the proposed model. Seitz et al. [15] proposed a better matched model to natural movement by processing events using event-driven updates Hao et al. [16] used the lattice gas model with parallel updating rules to study the one-way flow in the channel. Fu et al. [17] modified the static field to obtain the influence of individual preferences of export centers on crowd evacuation efficiency. Shi and wang [18] introduced game theory to simulate complex interactions between pedestrians in a pedestrian flow systems and the model shows that a smaller pedestrian density results in shorter average evacuation time. Bouzat and Kuperman [19] used the lattice gas scheme with pedestrians behavior to analyze evacuation of a single door room, and found that under certain conditions cooperators can escape more successfully. Guo et al. [20] proposed a potential on-site algorithm to establish pedestrian route selection behavior in an individual model based evacuation process with discrete space representation, and the developed model can reproduce more route selection patterns in the scene. Chen and

Li [21] introduced magnetic analogy to describe the repulsive interaction between pedestrians or pedestrians and obstacles, and conducted an experiment based on a force-driving cellular automata model.

However, most of the research on crowd evacuation models did not take the situation of personnel load into account. In the rare research on the model of evacuees with luggage [11], it is only analyzed that the increase in the proportion of luggage carriers will delay the evacuation process in the case of fewer people. But the obstruction impact of abandoning luggage on human movement was not analyzed in a space-intensive environment.

This paper considers the difference in the movement speed between people carrying luggage and people throwing luggage, the dependence and hindrance of luggage on movement. Then evacuation simulations was carried out under different luggage carrying ratios to obtain the regularity of evacuation in different situations, which mainly includes three category. The first one is that people carrying luggage will move with luggage. The second one is that people carrying luggage will move without luggage. The third one is that some of people carrying luggage will move with luggage.

2 Model

2.1 Evacuation Model of Evacuees with Luggage

- Cell division.

The evacuation room is composed of $L \times W$ cells, and each cell responding to 0.50 $\times$ 0.50 m. Each cell can be occupied at most one person or one piece of luggage.

- Evacuee attributes.

The evacuees are divided into 2 categories. They are people carrying luggage and people without luggage. The evacuees marked with numbers pn ($pn = 1, 2, 3 \cdots$) are randomly distributed in the evacuation space. Each evacuee in the model only carries one piece of luggage at most. The luggage is one—to—one correspondence with the evacuees and only moves along with the person. Same number bn, ($bn = 1, 2, 3 \cdots$) is given to the carrier and his/her luggage, and in the course of the movement, the pedestrian searches for his luggage by matching the number bn.

- Cell attributes.

Five kinds of cell states are defined: occupied by evacuees with luggage, occupied by evacuees without luggage, occupied by luggage carried, occupied by obstacles such as partition walls and chairs, and unoccupied.

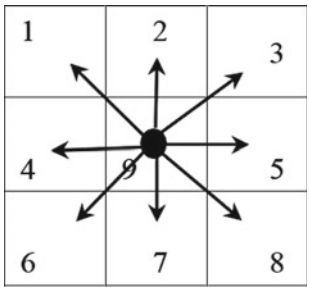


Fig. 1 Pedestrian movement direction

$$plaza_{xy} = \begin{cases} 0, & unoccupied \\ 1, & evacuees\ without\ luggage \\ 0.8, & evacuees\ with\ luggage \\ 0.4, & luggage \\ -0.95, & obstacles \end{cases} \quad (1)$$

- Movable direction.

Evacuees can move to empty neighborhood cells at some point in unit time. The possible directions are as shown in Fig. 1.

- Time step.

According to the study on the impact of luggage on the movement parameters of evacuees [1], the walking speed of those who do not carry luggage in an emergency is about 1.5 m/s, and is about 1.0 m/s of those carrying large luggage. Considering the slowing effect of luggage on the evacuation progress, and because all cellular status are updated synchronously according to the rules, those without luggage are updated every 2 time steps, while the carriers every 3 time steps.

2.2 Cellular Driven Model

- Distance parameter D

In order to leave the accident scene as soon as possible, evacuees will choose the closest direction to the exit. Since the evacuee can only move to neighboring cells at a time, the central cell calculates the distance between itself and its eight neighboring cells and the exit, respectively. The Euclidean distance to the nearest exit from the pedestrian's location is

$$d_{ij} = \min_n \sqrt{(i - i_n)^2 + (j - j_n)^2} \quad (2)$$

where (i, j) denotes the location of the cell chosen to update, (i_n, j_n) denotes location of the exit, and (x, y) denotes the location of cells in the evacuation space.

$$D_{xy} = \begin{cases} \frac{1}{d_{xy}}, \ x \in (i-1, i+1) \, and \, \mathrm{y} \in (j-1, j+1) \\ 0, \quad else \end{cases} \tag{3}$$

- Spatial parameter S

During the evacuation, people decide their next move position according to the current congestion situation. Due to the less resistance and more convenient movement, evacuees prefer to move to a larger space. Therefore, the central cell calculates the number of empty cells around itself and its eight neighbor cells s.

$$S_{xy} = \begin{cases} s_{xy}, \ x \in (i-1, i+1) \, and \, \mathrm{y} \in (j-1, j+1) \\ 0, \quad else \end{cases} \tag{4}$$

- Empty parameter B

In the process of moving, the evacuees will choose unoccupied cell as the target location, or choose to wait in place.

$$B_{xy} = \begin{cases} 0, \ plaza(x, y) \neq 0 \, and \, x \neq i, y \neq j \\ 1, \ plaza(x, y) = 0 \\ 1, \ plaza(x, y) = 0.8 \, or \, 1, \ \text{and} \, x = i, y = j \\ 1, \ plaza(i, j) = 0.8 \, and \, x = i_b, y = j_b \end{cases} \tag{5}$$

where (i_b, j_b) denotes the location of the luggage to be updated.

- Obstruction parameters R

In order to avoid falling into corners, the evacuees have a tendency to keep away from obstacles such as walls and seats, and the obstacle avoidance parameters are positively related to the distance from obstacles.

$$r_{ij} = \min_{n} \sqrt{(i - i_r)^2 + (j - j_r)^2} \tag{6}$$

$$R_{xy} = \begin{cases} r_{xy}, \ x \in (i-1, i+1).and.\mathrm{y} \in (j-1, j+1) \\ 0, \quad else \end{cases} \tag{7}$$

- Moving benefit value E

At each time step, the distance parameter is normalized according to Eq. (8). Space parameter and obstacle avoidance parameter are normalized according to Eq. (9), and the four parameters are weighted to obtain the moving benefit value.

$$\hat{X}_{\mathrm{xy}} = \frac{max\, X_d - X_{xy}}{max\, X_d - min\, X_d} \quad x \in (i-1, i+1) \text{且} y \in (j-1, \ j+1) \tag{8}$$

$$\hat{X}_{\mathrm{xy}} = \frac{X_{\mathrm{xy}} - \min X_d}{\max X_d - \min X_d} \quad x \in (i-1, i+1) \text{且} y \in (j-1,\ j+1) \tag{9}$$

$$E_{xy} = (\alpha \widehat{D_{\mathrm{xy}}} + \beta \widehat{S_{\mathrm{xy}}} + \delta \widehat{R_{xy}}) \cdot B_{\mathrm{xy}} \tag{10}$$

- Initial space occupancy M

The initial space occupancy density is the ratio of the total number of evacuees and luggage to the total number of spatial cells before the start of the evacuation, which is Eq. (12). m_a is the number of people without luggage. m_b is the number of people with luggage. m_c is the number of luggage.

$$M = \frac{m_a + m_b + m_c}{L \times W} \tag{12}$$

$$m_b = m_c \tag{13}$$

- Total number of evacuees in the system N

The total number of people in the initial system, including all people with luggage or not.

$$N = m_a + m_b \tag{14}$$

- The initial proportion of people with luggage K_1

The initial proportion of people with luggage is the ratio of the number of people carrying luggage to the total number of evacuees, before the start of the evacuation, which is

$$K_1 = \frac{m_b}{m_a + m_b} \tag{13}$$

- The proportion of discarded luggage K_2

The proportion of discarded luggage is the ratio of discarded luggage to total luggage, after the start of the evacuation, which is

$$K_2 = \frac{m_d}{m_c} \tag{14}$$

where m_d denotes the quantity of discarded luggage.

2.3 Evolutionary Rules

Through observation of the movement of people carrying luggage, it is found that the movement of the luggage relies on the movement of the owner, and thus the luggage and the person exist in a form of combination. In the process of moving forward, the owner's luggage is usually located on the side or behind, and the direction of movement is consistent with that of the person. In a crowded situation, the relative position of luggage and people will change, but still remain within a certain distance, otherwise people will lose control of the luggage, that is, if the luggage is abandoned, the luggage will not be able to move unless there are other people push it. Since the scattered luggage cannot move but still occupy space, its presence increases the possibility of pedestrians staying in place, and stop others movement. In order to avoid falling into dead angle, pedestrians will move the front obstacles in an emergency, even if the luggage doesn't belong to him. In this paper, the cell update rules are as follows:

- Pedestrian movement rules

  (1) For each update, the pedestrian selects the empty cell with the biggest moving benefit value as the target cell. If there are multiple cells in the neighborhood with the same maximum value, one of them is randomly selected as the target cell.
  (2) If multiple pedestrians compete for the same empty cell, the system randomly selects one of them to move to the cell in the next time step with the same probability, and the unselected stay in place.

- Luggage following rules

  (1) If a pedestrian moves with luggage, the luggage cannot be separated from the person and the distance between them cannot exceed $\sqrt{2}cell$. If a pedestrian abandons the luggage, the luggage will remain in place and will not move.
  (2) If the person corresponding to the luggage does not move, the luggage does not move.
  (3) If the owner can moves, the luggage is considered to move in the same direction so that the two move in parallel first. If the parallel movement rule cannot be satisfied, the luggage can be consider to move to the position where the person stay in at the previous time step.

- New route created Rules
  If people are surrounded by abandoned luggage, seats, walls, and other obstacles, the scattered luggage can be pushed forward to form a new advancing path. When pushing luggage, it is preferable to select the cell with the largest distance parameter.

3 Numerical Simulation and Result Analysis

3.1 Evacuation Simulation Model

In order to verify the validity of the proposed model, we take a train station waiting room as an example. A 20 × 20 m simulation system consisting of evacuation luggage, evacuees, and obstacles such as seats and walls is established using MATLAB. The speed of those carrying luggage is 1.0 m/s, and individual without luggage is about 1.0 m/s of those who do not carry luggage. Each time step of the simulation system is equivalent to 1/6 s, the system exit size is 3 m, and the exit pass speed is 1.0 m/s. In the simulation, the weight coefficient is taken as $\alpha = 0.5, \beta = 0.25, \delta = 0.25$, and the system evacuation time equals the total time from when the pedestrian begin to move toward the exit to the last person reached the exit. The simulation system will verify the model from three aspects: the effect of the evacuation process, the characteristics of expected speed, and the characteristics of the initial space occupancy. Based on this, three situations with different luggage carrying ratios: complete carrying of luggage, partial abandonment of luggage and complete abandonment of luggage are simulated to study the characteristics of evacuation with luggage.

3.2 Simulation Results Analysis

The layout of the system scene is shown in Fig. 2a. The simulation scene is a room with a single exit. Benches in the hall are fixed. At t = 0, people are seated on the seats as many as possible. Those without seats are randomly scattered in the system space (see Fig. 2b). At this time, people without luggage (blue), people carrying luggage (black) and the luggage (gray) occupies a space occupancy of 0.45, the initial proportion of people with luggage $K_2 = 0.5$. In the event of an emergency, people will quickly leave the seat and move to the exit in accordance with the rules of movement. Figure 2c shows the movement with all luggage carried and Fig. 2d shows that with all luggage abandoned. A typical phenomenon of evacuation will occur in the later period, that is, the crowd will form an arched jam at the exit. Figure 2e, g shows movement with luggage all-carried and all-abandoned in this situation respectively. Due to the randomness of the luggage abandonment and the influence of fixed seats on evacuation routes, the peak of the arched crowding phenomenon may be offset. From the self-organization phenomena such as jamming, clogging and oscillation at a bottleneck, it can be seen that the effect of the simulation evacuation process is in accordance with the actual process in a way.

When $M = 0.1, K_1 = 0$, it denotes that evacuees are all individuals without luggage. When $M = 0.2, K_1 = 1$, it denotes that evacuees are all people carrying luggage. Figure 3 shows that when there are 160 people in the system, the evacuation time will show the trend of decreasing first and then increasing with the increase of the expected speed, regardless of whether it is carrying baggage or individual

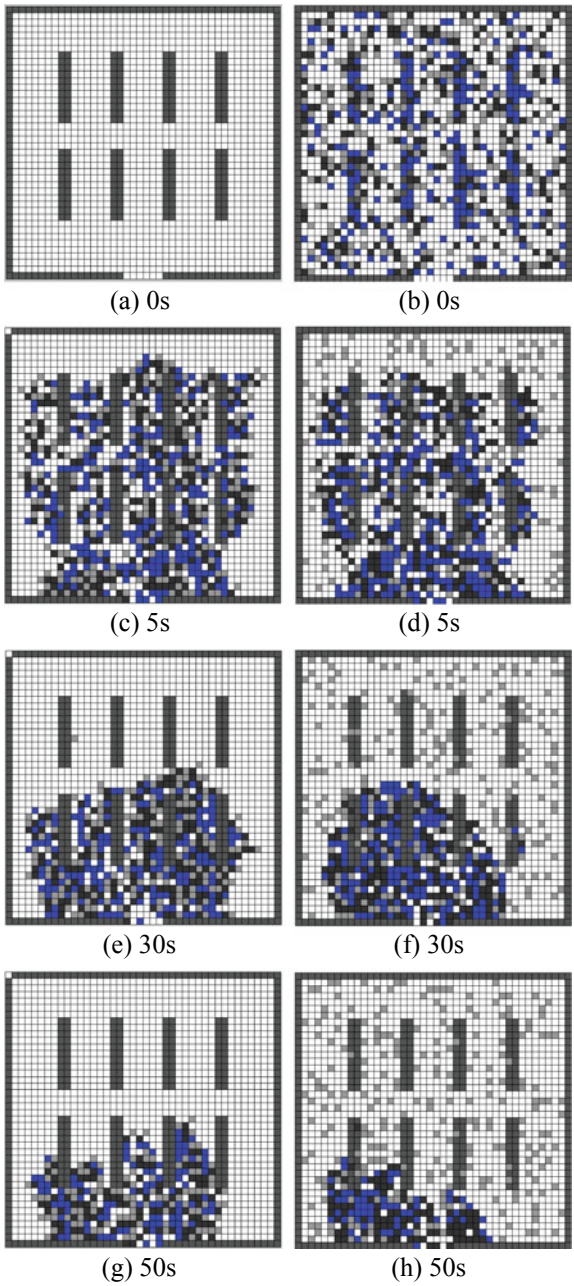


Fig. 2 Numerical simulation of evacuation at different time

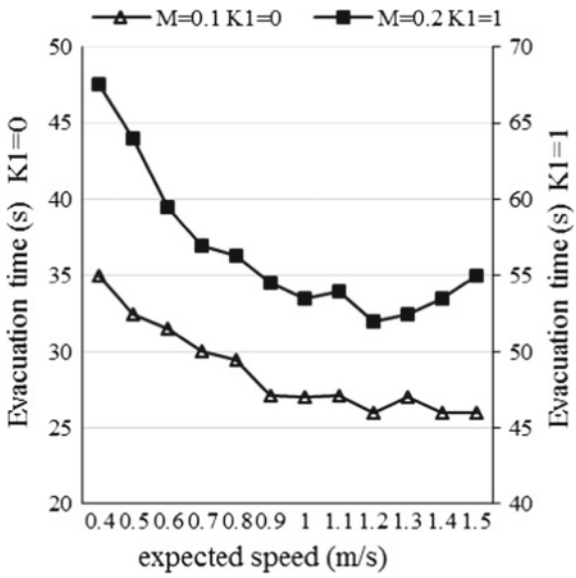


Fig. 3 Relation between the system evacuation time and desired velocity

movement. It is a typical phenomenon "fast is slow" in the pedestrian and evacuation dynamic (PED) models. In reality, since the maximum passing capacity at the exit is fixed, when the actual flow rate is less than the maximum capacity at the exit, increasing the desired speed can reduce the evacuation time; however, when the desired speed increases, the actual outlet flow rate will increase. When it is greater than the maximum flow rate at the exit, pedestrians will accumulate at the exit, congested with each other, increasing the probability of being trapped and increasing the evacuation time. It can be seen that the model simulation results are consistent with the actual situation.

Figure 4 shows that the system evacuation time increases with the increase of the initial space occupancy when evacuees move with all luggage carried. Because the greater the space occupancy, the lower the probability of pedestrians moving, the greater the chance of being forced to stay in place, and the longer the system evacuation time. In the actual situation, the conflict between pedestrian flows is more obvious with higher initial space occupancy density, thus, it is more likely to experience congestion and increase evacuation time. The results of the system are similar to the actual situation, which shows that the initial space occupancy is an important factor affecting the evacuation time. In addition, it can be seen from Fig. 4 that under the low space occupation density, the proportion of people carrying luggage has little impact on the evacuation time, while as occupancy density grows, the impact of the proportion is more pronounced, and it presents a positive correlation with the evacuation time. When the initial space occupancy is 0.7, under different initial proportion of people with luggage, there is a discrepancy about 36 s in evacuation time. The gap accounts for 20.69%.

Figure 5 depicts the relationship between the evacuation time and the proportion

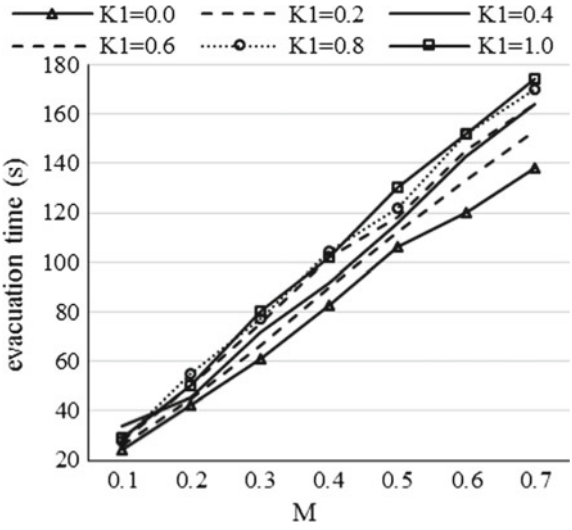


Fig. 4 Relation among the system evacuation time,initial space occupancy M, initial proportion of people with luggage K_1

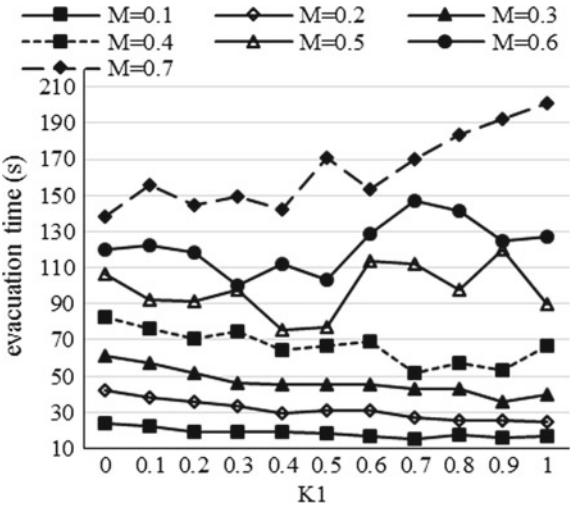


Fig. 5 evacuation of abandoned luggage

of people who initially carried luggage under the situation that luggage is completely discarded. When the initial space occupancy is less than or equal to 0.4, the evacuation time decreases with the increase of the proportion of the people who initially carried luggage. The reason is that under the same spatial density, the larger the proportion of initial carriers, the fewer the corresponding system evacuees, and thus the evacuation time is declining.

However, when the initial spatial density is greater than or equal to 0.4, the evacuation time of the system increases with the proportion, showing a trend of decreasing first and then changing. The reason is that when the proportion of people carrying

luggage is large, the amount of baggage scattered in the system space will increase, making the system space density slow down and the evacuation process decreases. The system evacuation time is mainly affected by the spatial density, so the evacuation time increases with the increase of abandoned luggage.

From the comparison chart of evacuation time between luggage all-carried and all-abandoned (as shown in Fig. 6), when the initial space occupancy density is less than or equal to 0.4, people give up carrying luggage to move. It significantly consumes less time than people carrying luggage, and the effect becomes more pronounced as the initial proportion of carried luggage increases. When the initial space occupancy is 0.3, the initial proportion of people with luggage is 1, the evacuation time gap between all carry luggage and all abnegate luggage can save time about 50.21%.

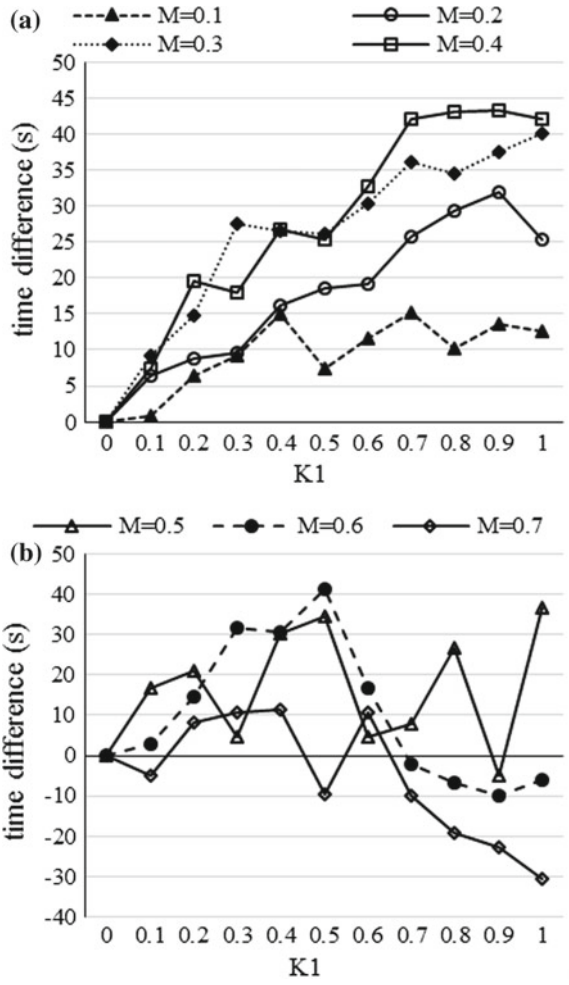


Fig. 6 **a** Carry-on luggage compared with carried luggage. **b** Carry-on luggage compared with abandoned luggage

However, as the initial space occupancy density increases to 0.6, especially when the proportion of people carrying luggage is greater than 50%, the evacuation time is shorter than that of carrying luggage. When the initial space occupancy is 0.7, carrying luggage rather than leaving luggage makes overall evacuation time shorter. Carrying luggage movement saves up to 31 s. The gap accounts for −17.62%, that is because when the initial space density is small, people who abandon luggage move faster than those carrying luggage, and the evacuation time is significantly reduced; but when the initial space density is large, the amount of luggage is large, people are crowded with each other, and scattered luggage obstructs the advancement channel. Thus, the speed advantage of giving up luggage is not come out in the case.

In order to study the impact of sectional abandoned luggage on the evacuation time, we set the same initial conditions, such as the same initial space occupancy, the same total number of evacuees, the same initial proportion of people with luggage. But they have different proportion of discarded luggage. In the part, all K_1 is 1. Each simulation system generates the experimental subject and luggage randomly, which induced the scene to be changeable. So the results may fluctuate. But we can still observe the regularity (as shown in Fig. 7). In the same initial personnel evacuation scale, when $M = 0.3$, initial space occupancy is a low lever, the evacuation time reduces as the proportion of people abandoning luggage increases, because, in this case, the number of scattered baggage is small, the force of obstruction is weak, and the speed get a raise. However, when $M = 0.6$, this trend has changed. In the process of evacuation, leaving luggage makes space occupancy cannot reduce timely, even stops the forward channel, so it may increase the evacuation time. We got an interesting discovery that blindly discarding luggage does not necessarily shorten the overall evacuation time of the system.

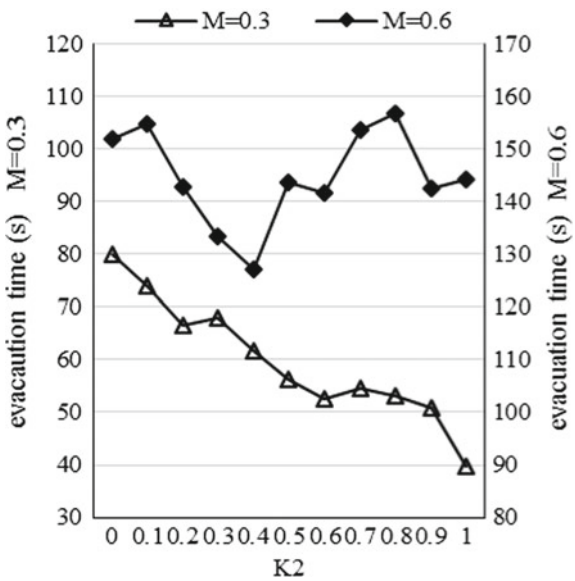


Fig. 7 Total evacuation time versus sectional abandoned luggage

4 Conclusion

In this paper, the cellular automata model is used to study the evacuation of people carrying baggage. Considering the rules of carrying luggage movement, different initial space occupancy and different proportions of initial luggage carriers are simulated to compare and analyze the influence of three types of behaviors, i.e. all luggage carried, all luggage discarded and part of luggage discarded on the evacuation time. Research indicates:

(1) The model quantifies the multiple parameters affecting pedestrian movement based on the characteristics of the luggage. In addition, the movement speed of people with and without luggage is distinguished. What is more, the model establishes the following rules of the luggage, and gives evacuees decision-making ability to get rid of the dead-end dilemma by new routes created rules. Besides, simulation scene is changed by the randomness of scattered luggage.
(2) This paper verifies the validity of the evacuation model from performance characteristics. This model successfully reproduce many self-organization collective phenomena, such as jamming and clogging, oscillation at a bottleneck and fast is slow.
(3) The simulation results show that when the luggage is fully carried, the increase of the initial space occupancy density makes the effect of increasing the proportion of luggage carried on the evacuation time more obvious. In addition, comparing the movements of carrying luggage and discarding luggage, it is concluded that when the initial space density is less than 0.4, the attachment of luggage plays a major role. The proportion of luggage carried increases, and the overall speed of the system slows down, as a result, the time needed for evacuation will increase. At this time, people should choose to abandon the luggage. When the initial space density is greater than 0.6, and the carried luggage ratio is greater than 0.5, the obstruction of the luggage will play a significant role. Discarding makes a large amount of the luggage become immovable obstacles and obstruct the advancement channel, which increase the evacuation time. Furthermore, with the same initial personnel size, study of the impact of partially discarding the luggage on the evacuation time was further verified that if the initial space density and the proportion of luggage carriers are not taken into account, blindly discarding luggage during the evacuation will not necessarily shorten the overall evacuation time of the system.
(4) The model reveals the inherent law between the evacuation effect and the proportion of luggage carried by evacuees, which can provide theoretical research support for the evaluation and optimization of pedestrian flow evacuation guidance strategy. Especially for places such as train stations, bus stations, subway stations, etc., where people are crowded and mostly loaded.
(5) During the actual evacuation process, the size and weight of the carried luggage, the physical condition of the evacuees, the familiarity with the site, the degree of panic about the emergency event, the layout of the system scene, the size of

the export, etc., may affect the evacuation time. Multiple mixed speed models under the influence of various factors are the goals of the next research.

References

1. Zhang K, Qiu Q, Wang R (2015) Study on effect of luggage on movement parameters of evacuees. China Saf Sci J 25(07):105–109
2. Li, X (2013) Research on personnel evacuation of high speed railway tunel fire. Master dissertation, Central South University, Changsha
3. Zhang L, Yue H, Li M et al (2015) Study on evacuation behaviors by a force-crowded simulation. Acta Phys 06:92–101
4. Feng S, Ding N, Chen T et al (2013) Simulation of pedestrian flow based on cellular automata: a case of pedestrian crossing street at section in China. Phys A 392(13):2847–2859
5. Varas A, Cornejo M, Mainemer D (2007) Cellular automaton model for evacuation process with obstacles. Phys A 382(2):631–642
6. Fu L, Zou N (2017) Cellular automata model of pedestrian flow at chool intersection. J Transp Syst Eng Inf Technol 03:98–104
7. Ren G, Ding C, Lu L et al (2014) Model and simulation of group pedestrian flow on a crosswalk based on cellular automaton. J Transp Syst Eng Inf Technol 14(2):45–50
8. Cheng Y, Zheng X (2018) Can cooperative behaviors promote evacuation efficiency? Phys A 492:2069–2078
9. Luo L, Fu Z, Cheng H et al (2018) Update schemes of multi-velocity floor field cellular automaton for pedestrian dynamics. Phys A 491:946–963
10. Fu Z, Zhou X, Zhu K et al (2015) A floor field cellular automaton for crowd evacuation considering different walking abilities. Phys A 4(20):294–303
11. You L, Zhang C, Hun J et al (2014) A three-dimensional cellular automaton evacuation model with dynamic variation of the exit width. Appl Phys 11(05):224–905
12. Hun J, You L, Zhang H et al (2018) Study on queuing behavior in pedestrian evacuation by extended cellular automaton model. Phys A 48(09):112–127
13. Chen C, Wang N, Xi B (2014) Research on evacuation cellular automaton model of evacuees with luggage. China Saf Sci J 24(7):3–9
14. Vermuyten H, Beliën J, De Boeck L et al (2016) A review of optimisation models for pedestrian evacuation and design problems. Saf Sci 87:167–178
15. Seitz M, Köster G (2014) How update schemes influence crowd simulations. J Stat Mech: Theory Exp 2014(7):P07002
16. Hao Q, Hu M, Cheng X et al (2010) Pedestrian flow in a lattice gas model with parallel update. Phys Rev E 82(2):026113
17. Fu Z, Yang L, Chen Y, Zhu K, Zhu S (2013) The effect of individual tendency on crowd evacuation efficiency under inhomogeneous exit attraction using a static field modified FFCA model. Phys A 392(23):6090–6099
18. Shi D, Wang B (2013) Evacuation of pedestrians from a single room by using snowdrift game theories. Phys Rev E 87:022802
19. Bouzat S, Kuperman M (2014) Game theory in models of pedestrian room evacuation. Phys Rev E 89:032806
20. Guo R, Huang H, Wang S (2013) A potential field approach to the modeling of route choice in pedestrian evacuation. J Stat Mech: Theory Exp 02:P02010
21. Chen C, Li J, Zhang D (2012) Study on evacuation behaviors at a T-shaped intersection by a force-driving cellular automata model. Phys A 391(07):2408–2420

Experiment on Destination Choice Game

Haoran Li and Chuanci Cai

Abstract A fundamental issue in traffic science is to understand the behavior of traffic participants. Ether to theoretically model the traffic phenomenon or to analyze the complex operation process of the actual traffic system, the underlying mechanism of individual traffic decision-making behavior has always been the focus of relevant scholars. Many researches on traffic decision-making have been done, but the research on destination choice which is the most basic motivation for individual travel is still lacking. In this paper, we establish a simplified laboratory experiment to study individuals' destination choice behaviors. Considering that increasingly diverse traffic information is provided to residents through the Internet, two treatments are set up to explore the influence of feedback information on human behaviors. In a lot of real scenarios, the individual is most likely to take the degree of congestion into account, so we set the payoff of each destination to a form negatively related to that degree. Experimental results demonstrate that aggregate behavior would achieve user equilibrium rapidly no matter whether there is feedback information or not. Moreover, the feedback information has a certain effect on individual choice behaviors, which is a socially significant discovery. We believe that the results can be combined with other models for comprehensive modeling, which will be helpful to the traffic planning and management in the future.

Keywords Laboratory experiments · Destination choice · Human behaviors · Feedback information

1 Introduction

How to better describe and analyze human behaviors has always been a common concern in many fields such as biology, psychology, economy and traffic. However,

H. Li (✉) · C. Cai
Institute of Transportation System Science and Engineering, Beijing Jiaotong University, Beijing 100044, China
e-mail: 16120752@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_43

in the field of traffic, it has always been a great challenge to analyze traffic macroscopic or microscopic phenomena from the perspective of human choice behaviors. The key to study this kind of problem is to get real and effective individual decision data. In traditional traffic research, traffic survey methods (mainly including revealed preference and stated preference) [1] are usually used to collect the data that reflect individuals' travel decision-making behaviors. However, these methods have certain limitations, through which neither can we restore the real competition scenarios, nor control the key factors that affect individual travel decisions, let alone observe the impact of unimplemented traffic management strategies. In recent years, many scholars have used the methods of laboratory experiments in economics to abstract travel decision-making problems in traffic into non-cooperative crowded game experiments. By using certain material stimulation to simulate the real decision-making environment, they successfully verify the network balance problem and the classic paradoxes [2] in traffic, in addition, some researchers completed some evaluations of management measures related to traffic demand that have not yet been implemented.

On the problem of travel decision-making, scholars mainly focused on the individuals' behaviors of route choice [3–8], departure time choice [3, 9] or travel mode choice [10] and a considerable number of achievements have been achieved. However, these studies ignored the destination choice which is the most fundamental motivation of individual travel.

In this paper, we use the idea of experimental economics to study individual destination choice behaviors by conducting 8 groups of experiments (4 groups for Local and 4 groups for Global) under different treatments in which the treatment Local offers subjects' own historical information and Global offers all subjects' historical information who belong to the same system aiming to explore whether feedback information has an impact on the system and subjects. Figure 1 shows the major differences of two treatments on the user interfaces. Through the analysis of the experimental data, it can be concluded that all the groups under different treatments rapidly achieve the user equilibrium [11], and keep fluctuating around the equilibrium points afterwards. Feedback information has no obvious effect on the process to achieve the user equilibrium but it can affect individual decision-making.

2 Experimental Set-up

There are one origin and four destinations A, B, C, and D in our experiments as shown in Fig. 2 and it is assumed that all these destinations can satisfy certain needs of all the subjects in the system. The total rounds of each group of experiment is set to 200, and the subjects were kept confidential. In each round of experiments, one has to choose a destination within a given time (20 s) and receive the corresponding payoff which will decrease with the increase of the number of people who choose the same destination. Different destination corresponds to different descent rate. The payoff corresponding to the destination i is calculated by Eq. (1) and the true reward of each subject is positively related to the accumulated payoff.

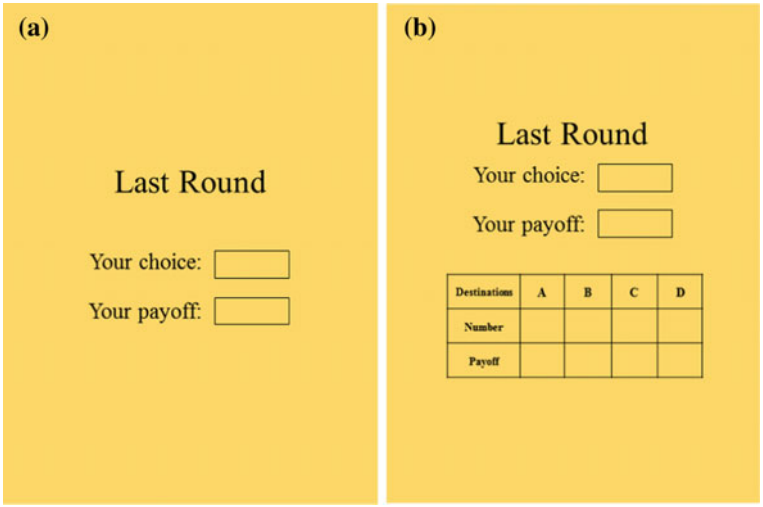


Fig. 1 Partial screenshots of user interfaces. **a** Treatment local. **b** Treatment global

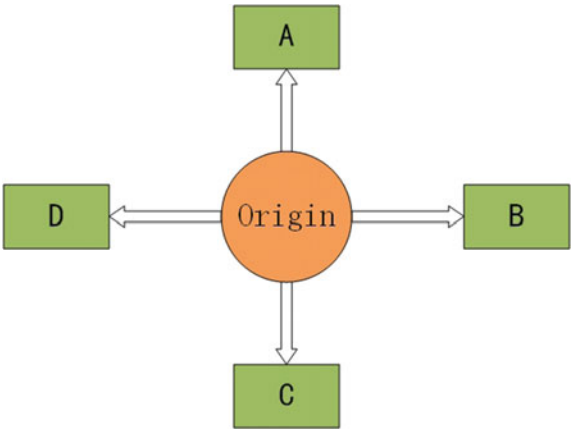


Fig. 2 Schematic of experimental traffic network

$$Payoff_i = 50 - k_i \times n_i \tag{1}$$

where k_i is the descent rate of destination i (in our experiments, we set the value 2, 4, 6 and 8 for the destinations, respectively) and n_i represents the number of people choosing the destination i.

The number of subjects in each group was 25, mostly postgraduates and undergraduates from Beijing Jiaotong University. 8 groups (4 for the treatment Local and 4 for the treatment Global) of experiments were conducted.

The theoretically user equilibrium under two treatments which is shown in Table (1) can be calculated by Eq. (2) easily.

Table 1 Theoretically user equilibrium under two treatments

Treatments	Items	k_1	k_2	k_3	k_4
Local/Global	Value	2	4	6	8
	Population	12	6	4	3
	Payoff (UE)	26	26	26	26

$$\min : \mathrm{Z(X)} = \sum_j \int_0^{T_{ij}} t_j(\omega) d\omega$$
$$\text{s.t.} \begin{cases} \sum_j T_{ij} = m \\ T_{ij} \geq 0 \end{cases} \tag{2}$$

3 Aggregate Behaviors

User Equilibrium (UE) is a profile of strategies such that individuals independently minimize their private travel costs and do not have a tendency to change their choices. Regarding to our experiments, the theoretical UE can be obtained from Eq. (2), i.e., the costs of all destinations are equal. The mean number of commuters at each destination in our experiments are shown in Fig. 3. The figures show that the mean

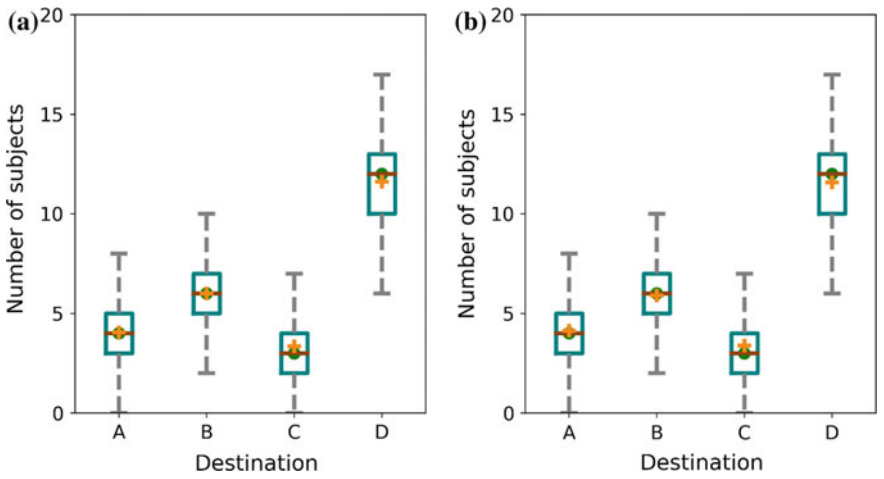


Fig. 3 Aggregately comparing UE and mean or median values of number of subjects. **a** Treatment local. **b** Treatment global. The green dots denote UE. The brown shot lines denote median values of number of subjects. The orange crosses denote mean values of number of subjects. The boxplots denote interquartile ranges of number of subjects

and median numbers of subjects are quite close to the theoretical UE. Moreover, we find there is no obvious difference for the mean and median numbers of subjects at each destinations between two different information feedback treatments.

Figure 4 shows that the number of subjects in each group can converge to theoretical UE in about 15 rounds and the persistently fluctuates around it over the entire period in treatment Local and treatment Global. The feedback information has no

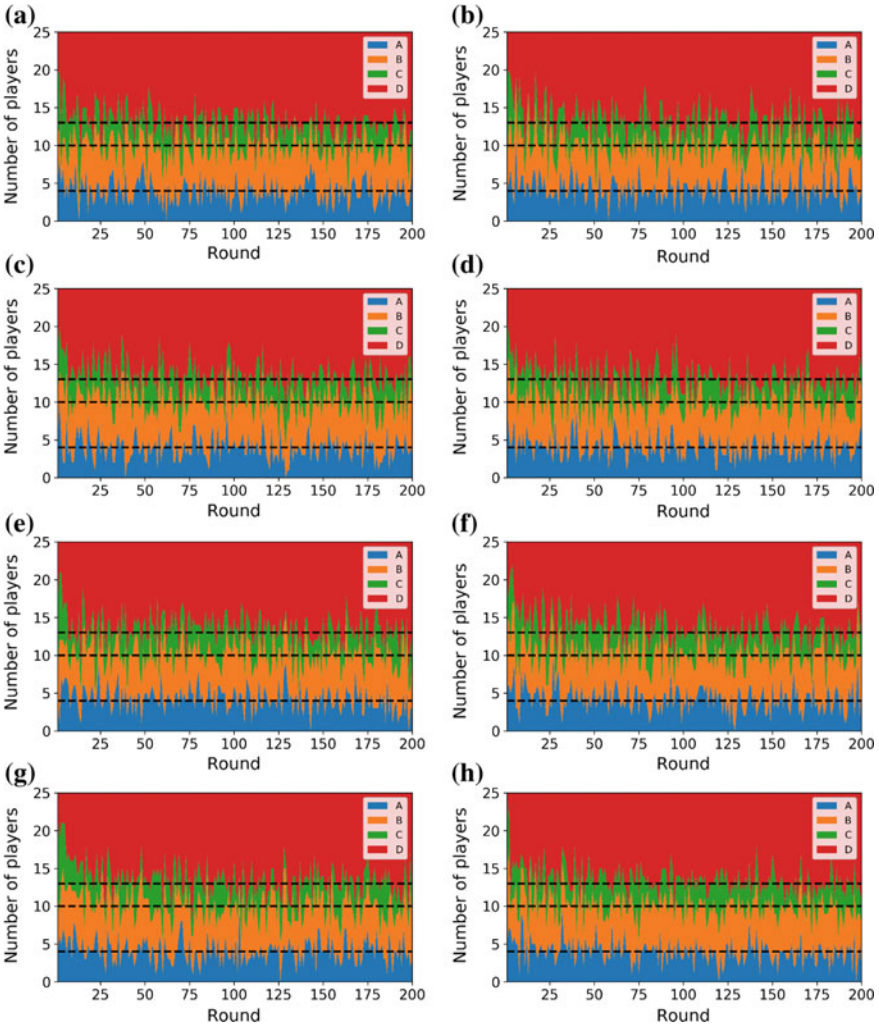


Fig. 4 Time evolution of number of subjects in the experiments. Stacked plots of the observed number of subjects at each destinations under treatment Local [left panel, including (**a**), (**c**), (**e**), (**g**)] and under treatment Global [right panel, including (**b**), (**d**), (**f**), (**h**)]. The horizontal dash lines denote UE

significant impact on the process to achieve UE, and the self-information of subjects (Local) is sufficient to facilitate the achievement of UE.

4 Individual Behaviors

4.1 Destination Changes

As feedback information has no evident impact on the process of achieving UE, we then explore whether subjects' behavior is affected by the information. Figure 5 shows the aggregate mean value of destination change rates [12] as a function of rounds. Apparently there is a negative trend in treatment Global and the change rate has dropped from the very first 0.68 to around 0.4. As for another treatment the trend is more stable and the change rate has been fluctuating around 0.5 throughout the whole period although there is also a decrease during the first dozens of rounds.

From the perspective of collectivity, we divided the entire experimental process into 3 segments based on Fig. 5. The first one contains the first 50 rounds. In this segment, the change rates of both treatments have declined to some extent and it can be easily observed that the change rates of the treatment Global are higher than those of the Local (a two-sided Mann-Whitney U-test rejects the null hypothesis, $p < 0.05$), which can be understood that in the early period subjects familiarize themselves with different strategies by changing strategies and the information feedback has the effect of prompting subjects to change strategies more frequently. The second one contains the next 100 rounds. Compared to the previous segment, the change rates of both treatments drop to around 0.5 and it is difficult to find the difference between Global and Local (a two-sided Mann-Whitney U-test accepts the null hypothesis), which suggested in this segment of both treatments the subjects will not change strategies

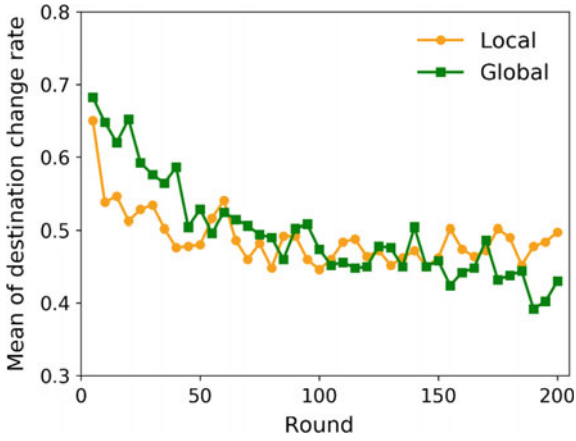


Fig. 5 Aggretate time evolution of change rates under different treatments. The orange dots and polygonal line denote destination change rates and trend under treatment Local. The green squares and polygonal line denote destination change rates and trend under treatment Global. Each circle (square) represents the mean value of destination change rates of 4 groups under treatment Local (Global) in bins of 5 rounds

frequently as before and begin to carefully consider whether frequently changing strategies will result in improved earnings. The third one contains the last 50 rounds, in which the rate of the treatment Local is stable at around 0.5 while the rate of the treatment Global keeps a downward trend to around 0.4 (a two-sided Mann-Whitney U-test rejects the null hypothesis, $p < 0.05$). It is suggested that the subjects in the treatment Global have realized that frequent changing strategies have no significant trend to increase payoffs and apparently they realized this earlier than those in another treatment.

4.2 Response Modes

In the previous section, we used to analyze from the perspective of groups. Now we are digging deeper into individual behavior from an individual perspective. In this part, we use Yule coefficient [6] considering the subjects who change his/her strategies to recognize the response mode of subjects. Specifically, a subject is classified to be in direct response mode if he/she changes (does not change) his/her strategy when his/her payoff is below (exceeds) the UE payoff (calculated in advance, 26) or contrary response mode if he/she changes (does not change) his/her strategy when his/her payoff exceeds (is below) the UE payoff. And the Yule coefficient for each subject can be calculated by

$$\mathrm{Q} = \frac{C_- \times S_+ - C_+ \times S_-}{C_- \times S_+ + C_+ \times S_-} \tag{3}$$

where C_- and C_+ are the numbers of times a subject changes his/her strategy (C_- for a payoff below 26 and C_+ for a payoff above 26), S_- and S_+ are the numbers of times a subject keeps his/her strategy unchanged (S_- for a payoff below 26 and S_+ for a payoff above 26). The Yule coefficient which is marked as Q ranges from -1 to $+1$.

In our case, a high Yule coefficient reflects a tendency towards direct responses and a low one a tendency towards contrary responses. Here we classifies subjects with Q above $+0.5$ as direct responders, subjects with Q below -0.5 as contrary responders. Classification of response patterns, mean value of Q and corresponding standard deviation for each group are shown in Table 2. It can be clearly observed that the mean values of Q for treatment Local are higher than Global and the proportion of the direct responders for treatment Local is obviously higher than Global, which suggested that under the condition of poor information, subjects are more inclined to direct response.

Table 2 Statistical results of the Yule coefficients Q in the experiments. The percentage of $0.5 < Q < 1$ and mean value of Q under treatment Local are larger than that under treatment Global

Treatments	Groups	$-1 < Q < -0.5$	$0.5 < Q < 1$	$-0.5 \leq Q \leq 0.5$	Mean(Q)	Std(Q)
Local	1	0.08	0.68	0.24	0.5039	0.5454
	2	0.24	0.44	0.32	0.2876	0.6902
	3	0.16	0.32	0.52	0.2222	0.5679
	4	0.08	0.36	0.56	0.2634	0.5069
	Mean	0.14	0.45	0.41	0.3193	
Global	5	0.04	0.12	0.84	0.1705	0.3243
	6	0.04	0.24	0.72	0.0644	0.5164
	7	0	0.2	0.8	0.3326	0.4395
	8	0.08	0.48	0.44	0.2004	0.3810
	Mean	0.04	0.26	0.7	0.1920	

5 Conclusion and Discussion

Due to the advantages of experiment it can offer in controllability and repeatability and restore real competition scenario through material stimulation, experimental methods are becoming more and more popular for traffic researchers. In recent years, researches on travel choice behavior experiment mainly focused on the choice of departure time, route selection and travel mode selection. Learning from their experience, in this paper, we have experimentally explored the human destination choice behavior which is the initial motivation of human mobility and almost never been studied.

Experimental results show that the collective behavior approaches theoretically UE, regardless of different information feedback treatments. In all experimental groups, we observed that the number of people in each destination converged to the equilibrium quickly in both treatments, and fluctuations existed around UE throughout the whole process which shows that the historical information has no significant effect on the process of convergence to UE.

At the individual level, we have identified the impact of different treatments on human behavior. We found that global historical feedback information would prompt subjects to reduce the frequency of strategies changing more quickly by calculating subjects' strategies changing rate in different treatments, which means that if such traffic information is provided in reality, it would reduce the frequency of residents' changing destinations. For a long time, to some extent, providing such information may make more rational use of social resources and promote the stability of social order. By calculating Yule coefficient, two response mode can be found in the data, a direct one in which road changes follow bad payoffs and a contrary one in which road changes follow good payoffs. Comparing Yule coefficients of different treatments, we found that subjects in treatment Local are more inclined to direct response, in

other words, the global feedback information can really help residents to think more comprehensively and not to rush to make decisions.

References

1. Hensher DA (1994) Stated preference analysis of travel choices: the state of practice. Transportation 21(2):107–133
2. Mak V, Seale DA, Gisches EJ, Yang R, Cheng M, Moon M et al (2017) The braess paradox and coordination failure in directed networks with mixed externalities. Prod Oper Manag
3. Mahmassani HS, Chang G-L, Herman R (1986) Individual decisions and collective effects in a simulated traffic system. Transp Sci 20:258–271
4. Mahmassani HS, Stephan D (1988) Experimental investigation of route and departure time dynamics of urban commuters. Transp Res Rec 1203:69–84
5. Helbing D, Schönhof M, Kern D (2002) Volatile decision dynamics: experiments, stochastic description, intermittency control and traffic optimization. New J Phys 4:33.1–33.16
6. Selten R, Chmura T, Pitz T, Kube S, Schreckenberg M (2007) Commuters route choice behaviour. Games Econ Behav 58:394–406
7. Zhao CL, Huang HJ (2014) Experiment of boundedly rational route choice behavior and the model under satisficing rule. In: IEEE 17th international conference on intelligent transportation systems (ITSC) October 8–11, Qingdao, China
8. Rapoport A, Gisches EJ, Mak V (2014) Distributed decisions in networks: laboratory study of routing splittable flow. Prod Oper Manag 23(2):314–331
9. Sun X, Han X, Bao JZ, Jiang R, Jia B, Yan X et al (2017) Decision dynamics of departure times: experiments and modeling. Phys A 483:74–82
10. Liu C, Mak V, Rapoport A (2015) Cost-sharing in directed networks: experimental study of equilibrium choice and system dynamics. J Oper Manag 39–40:31–47
11. Wardrop JG (1952) Some theoretical aspects of road traffic research. Proc Inst Civil Eng Part II 1:325–378
12. Mak V, Gisches EJ, Rapoport A (2015). Route versus segment: an experiment on real-time travel information in congestible networks. Prod Oper Manag, 24(6):947–960

Alleviate Traffic Congestion and Reduce Energy Consumption by Setting a Peak-Only Bus Lane on a Bottleneck-Constrained Highway

Xingfei Wang and Xingang Li

Abstract With the great popularity of the public transit which is a kind of green transportation, peak-only bus lane is gradually implemented on the corridors of large cities to make the bus runs a privilege to go through the bottleneck, then the bus runs can keep a faster speed which will definitely attract more potential commuters. And thus this will alleviate the traffic congestion caused by private cars and reduce the energy consumption and emissions to some degree. In this paper, we investigate the impact of the peak-only bus lane on alleviating traffic congestion and reducing energy consumption by using bottleneck model with auto and bus modes. The peak-only bus lane will occupy the bottleneck's capacity by a fixed amount just within a fixed peak hour period. While the capacity of the bottleneck for auto mode commuters is time-varying within the whole commuting period. It is assumed that the mode choice and the departure time choice are governed by the user equilibrium criterion and nobody can decrease his/her total cost by adjusting the mode or the departure time in the equilibrium state. The departure rates for both bus and auto modes are derived analytically. The travel cost and energy consumption are analyzed with different bus dispatch frequencies and bus lane capacities. The numerical results showed that the setting of the peak-only bus lane will descend the number of commuters who choose the auto mode, and thus decrease the system's total travel cost and energy consumption to certain degree. The optimal setting of the road resources and the frequency for peak-only bus lane was also investigated. We believe that the results are helpful to the planning and operating of peak-only bus lane, and it's useful to alleviate traffic congestion, reduce the energy consumption and protect the environment.

Keywords Bottleneck model · Peak-only bus lane · User equilibrium · Energy consumption

X. Wang · X. Li (✉)
Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China
e-mail: lixingang@bjtu.edu.cn

X. Wang
e-mail: 16120754@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_44

1 Introduction

Vickrey [1] first applied the deterministic queuing theory in transportation and proposed a basic endogenous model that can lead all travelers have the same traffic cost in departure time choice, which was famous known as bottleneck model. A series of authoritative papers on this topic were continually published and kept up-to-date. In the absence of other travel modes, there are many advanced theories such as user equilibrium, dynamic equilibrium, deterministic queuing, road pricing, time-varying road toll, elastic demand, user heterogeneity, etc. There are many big contributions in the bottleneck model when adding a bus mode competing with the private cars.

Most of the existing literatures for bottleneck model with bus and auto modes, however, are based on static settings with either dedicated bus lane or shared lanes. In fact, dedicated lane has been well discussed in the previous studies, and we focus the shared lanes, i.e., peak-only bus lane. Most researches about bottleneck model when assuming capacity a constant over time within day. In the subsequent research, works on the bottleneck capacity has been extensive studied [2–4]. Moreover, there are few paper invested the energy consumption when setting a bus lane. It's essential to explore the effect of time-varying capacity to modal split, commuting pattern and the energy consumption.

This paper is to extend the model proposed by Huang [5] and investigate the effect of commuters' commuting behaviors and the energy consumption by vehicles on a bottleneck-constrained highway when there is a peak-only bus lane. User Equilibrium model are used to govern the mode choice. Then we present numerical results to analysis the effects when the peak-only bus lane is set or not from simulation. In the end, we concludes the paper according the numerical results. Notations used in this paper are listed in Table 1.

2 Model Analysis

2.1 Auto Mode

Consider a continuum of N commuters from home (H) to workplace (W) via a corridor with more than two lanes (one peak-only bus lane) every morning. In this paper, all commuters are assumed to be identical in value of time and schedule delay penalty. There is a bottleneck (B) at the end of the highway, with a fixed capacity of S *CEQs* per unit time. Following the user equilibrium principle, the first and the last commuters meet same travel cost and no waiting time, i.e.,

$$\alpha T_1 + \beta\left(t_w - t_c^e\right) + p_1 = \alpha T_1 + \gamma\left(t_c^l - t_w\right) + p_1 \tag{1}$$

Without loss of generality, we assume that the operation period for peak-only bus lane is within the car queuing period, which means that $t_c^e < t_z^e < t_z^l < t_c^l$. Since the

Table 1 Notations adopted in the proposed model

Symbol	Definition	Remarks
Given parameters		
N	Total numbers of the commuters every day	
α, β, γ	The value of unit travel time, the shadow cost of early and late	$\gamma > \alpha > \beta$
t_w	Official work starting time (common preferred)	
P_1, P_2	The fixed costs of car and bus mode	
T_1, T_2	The fixed time of car and bus from H to the bottleneck	
S, S_b	Capacity of the bottleneck and resources occupied by peak-only bus lane	unit *CEQ*
λ	One bus equal to λ *CEQ*s (one car equal to 1 *CEQ*)	
f	The frequency of dispatching buses	$\lambda f < s$
t_z^e, t_z^l	The start time and end time of the peak-only bus lane	$t_z^e < t_w < t_z^l$
Decision variables		
N_1, N_2	The numbers of the auto and bus mode commuters	$N_1 + N_2 = N$
$\hat{t}_c, \hat{t}_b$	The time of arriving at bottleneck of auto mode and bus mode commuters who can arrive at workplace on time respectively	
t_c^e, t_c^l	The start and end time of the queue formed by cars	
t_b^e, t_b^l	The start and end time of the bus runs join the queue with the cars	
$\hat{t}_z^e, \hat{t}_z^l$	The time of arriving at bottleneck so the auto mode commuter can arrive at workplace at t_z^e and t_z^l rightly	
$\delta_c(t)$	The schedule delay cost of auto mode commuters who arrive at bottleneck at t	
$C_1(t)$	The total travel cost of an auto commuter departing at t	unit *HK* $
$T_{w\_c}(t)$	The waiting time in the queue of an auto commuter departing at t	
$r_c(t)$	The departure rate of auto commuters at time t	unit cars/2 min
$C_2(i)$	The total cost of a commuter who chose the bus mode on the i th bus run	unit *HK* $
t_i	The time that the i th bus run arrive at the bottleneck	
n_i	The number of the commuters who chose the i th bus run	
$T_{w\_b}(t_i)$	The waiting time in the queue of the bus run who arrives at bottleneck at t_i	
$\delta_b(t_i)$	The schedule delay cost of a bus commuter who arrives at bottleneck at t_i	
ω	There are ω batches bus runs which transport at least one commuter	
cong(x)	The in-carriage congestion function of the bus mode commuters	
EC	Total energy consumption (unit Con)	
EC_c	The energy consumption of each private car (unit Con)	
$EC_b(n_i)$	The energy consumption of a bus runs with n_i passengers (unit Con)	

bottleneck is operated in full capacity from t_c^e to t_c^l, one can have the equation,

$$(S-\lambda f)\left(t_z^e-t_c^e\right)+(S-S_b)\left(t_z^l-t_z^e\right)+(S-\lambda f)\left(t_c^l-t_z^l\right)=N_1 \tag{2}$$

With Eqs. (1) and (2), we can derive the time the queue first forms t_c^e and totally dissipates t_c^l,

$$t_c^e=t_w-\frac{\gamma\left[N_1-\left(t_z^l-t_z^e\right)(\lambda f-S_b)\right]}{(S-\lambda f)(\beta+\gamma)} \tag{3}$$

$$t_c^l=t_w+\frac{\beta\left[N_1-\left(t_z^l-t_z^e\right)(\lambda f-S_b)\right]}{(S-\lambda f)(\beta+\gamma)} \tag{4}$$

In the equilibrium, commuters face same travel cost, then one can get the equations:

$$\alpha\left[T_1+\left(t_w-\hat{t}_c\right)\right]+p_1=\alpha T_1+\beta\left(t_w-t_c^e\right)+p_1 \tag{5}$$

$$\alpha\left[T_1+\left(t_z^e-\hat{t}_z^e\right)\right]+\beta\left(t_w-t_z^e\right)+p_1=\alpha T_1+\beta\left(t_w-t_c^e\right)+p_1 \tag{6}$$

$$\alpha\left[T_1+\left(t_z^l-\hat{t}_z^l\right)\right]+\gamma\left(t_z^l-t_w\right)+p_1=\alpha T_1+\gamma\left(t_c^l-t_w\right)+p_1 \tag{7}$$

From the equations above, we can conclude that

$$\hat{t}_c=\frac{(\alpha-\beta)}{\alpha}t_w+\frac{\beta}{\alpha}t_c^e=t_w-\frac{\beta\gamma}{\alpha}\left(\frac{N_1-\left(t_z^l-t_z^e\right)(\lambda f-S_b)}{(S-\lambda f)(\beta+\gamma)}\right) \tag{8}$$

$$\hat{t}_z^e=\frac{(\alpha-\beta)}{\alpha}t_z^e+\frac{\beta}{\alpha}t_c^e=\frac{(\alpha-\beta)}{\alpha}\left(t_z^e-t_w\right)+\hat{t}_c \tag{9}$$

$$\hat{t}_z^l=\frac{(\alpha+\gamma)}{\alpha}t_z^l-\frac{\gamma}{\alpha}t_c^l=\frac{(\alpha+\gamma)}{\alpha}\left(t_z^l-t_w\right)+\hat{t}_c \tag{10}$$

Obviously, we can derive that $\hat{t}_z^e<\hat{t}_c<t_z^l$ from the Eqs. (8)–(10), which can be demonstrated by the rule (first in first out) that overtaking is forbidden. Next with these parameters determined, we can denote the equilibrium travel cost experienced by the auto commuters during $\left[t_c^e,t_c^l\right]$,

$$C_1=\alpha\left(T_1+T_{w\_c}(t)\right)+\delta_c(t)+P_1=\alpha T_1+\beta\gamma\xi+P_1 \tag{11}$$

Here $\xi=\frac{N_1-\left(t_z^l-t_z^e\right)(\lambda f-S_b)}{(S-\lambda f)(\beta+\gamma)}$. As we all know, the schedule delay cost contains two phases,

$$\delta_c(t) = \begin{cases} \beta\left[t_w - t - T_{w\_c}(t)\right], & t \le \hat{t}_c \\ \gamma\left[T_{w\_c}(t) + t - t_w\right], & t \ge \hat{t}_c \end{cases} \tag{12}$$

Combining Eqs. (11) and (12), and the waiting time in the queue for auto commuters who arrive at B at time t is

$$T_{w\_c}(t) = \begin{cases} \frac{\beta\left(t - t_c^e\right)}{\alpha - \beta}, & t \in \left[t_c^e, \hat{t}_c\right] \\ \frac{\gamma\left(t_c^l - t\right)}{\alpha + \gamma}, & t \in \left[\hat{t}_c, t_c^l\right] \\ 0, & otherwise \end{cases} \tag{13}$$

Considering the constant frequency, the arrival rate of bus mode at B is λf. Occupying some of the road resources, the setting of the peak-only bus lane will change the capacity for the private cars and the capacity will be time-varying for the auto mode. Zhang et al. [6] analyzed the bottlenecks with time-varying capacities when there is only auto mode. According to the analysis, the cumulative arrival curve to B is also a piecewise linear curve. The separation points are determined by t_c^e, $\hat{t}_z^e$, t_z^e, $\hat{t}_c$, $\hat{t}_z^l$, t_z^l and t_c^l. With the assumptions and calculated results, one can get the relationship that $t_c^e < \hat{t}_z^e < \left(t_z^e, \hat{t}_c < \hat{t}_z^l\right) < \hat{t}_z^l < \hat{t}_c^l$. Next the relationship of t_z^e, $\hat{t}_c$ and $\hat{t}_z^l$ will be discussed and the arrival rate at B will be given.

Case (i). $\hat{t}_c < t_z^e < \hat{t}_z^l$. Then one can get the arrival rate of the auto mode $r_c(t)$,

$$r_c(t) = \begin{cases} \frac{\alpha s}{\alpha - \beta} - \lambda f, & t \in \left[t_c^e, \hat{t}_z^e\right] \\ \frac{\alpha(s - s_b)}{\alpha - \beta} - \lambda f, & t \in \left[\hat{t}_z^e, \hat{t}_c\right] \\ \frac{\alpha(s - s_b)}{\alpha + \gamma} - \lambda f, & t \in \left[\hat{t}_c, t_z^e\right] \\ \frac{\alpha(s - s_b)}{\alpha + \gamma}, & t \in \left[t_z^e, \hat{t}_z^l\right] \\ \frac{\alpha}{\alpha + \gamma} S, & t \in \left[\hat{t}_z^l, t_z^l\right] \\ \frac{\alpha s}{\alpha + \gamma} - \lambda f, & t \in \left[t_z^l, t_c^l\right] \end{cases} \tag{14}$$

Case (ii). $\hat{t}_c < \hat{t}_z^l < t_z^e$. Then the relationship is $t_c^e < \hat{t}_z^e < \hat{t}_c < \hat{t}_z^l < t_z^e < t_z^l < t_c^l$ and

$$r_c(t) = \begin{cases} \frac{\alpha s}{\alpha - \beta} - \lambda f, & t \in \left[t_c^e, \hat{t}_z^e\right] \\ \frac{\alpha(s - s_b)}{\alpha - \beta} - \lambda f, & t \in \left[\hat{t}_z^e, \hat{t}_c\right] \\ \frac{\alpha(s - s_b)}{\alpha + \gamma} - \lambda f, & t \in \left[\hat{t}_c, \hat{t}_z^l\right] \\ \frac{\alpha s}{\alpha + \gamma} - \lambda f, & t \in \left[\hat{t}_z^l, t_z^e\right] \\ \frac{\alpha}{\alpha + \gamma} s, & t \in \left[t_z^e, t_z^l\right] \\ \frac{\alpha s}{\alpha + \gamma} - \lambda f, & t \in \left[t_z^l, t_c^l\right] \end{cases} \tag{15}$$

Case (iii). $t_z^e < \hat{t}_c < \hat{t}_z^l$, and one can obtain the arrival rate of the auto mode,

$$r_c(t) = \begin{cases} \frac{\alpha s}{\alpha-\beta} - \lambda f, & t\epsilon\left[t_c^e, \hat{t}_z^e\right] \\ \frac{\alpha(s-s_b)}{\alpha-\beta} - \lambda f, & t\epsilon\left[\hat{t}_z^e, t_z^{\mathrm{e}}\right] \\ \frac{\alpha(s-s_b)}{\alpha-\beta}, & t\epsilon\left[t_z^e, \hat{t}_c\right] \\ \frac{\alpha(s-s_b)}{\alpha+\gamma}, & t\epsilon\left[\hat{t}_c, \hat{t}_z^l\right] \\ \frac{\alpha}{\alpha+\gamma}s, & t\epsilon\left[\hat{t}_z^l, t_z^l\right] \\ \frac{\alpha s}{\alpha+\gamma} - \lambda f, & t\epsilon\left[t_z^l, t_c^l\right] \end{cases} \tag{16}$$

However, the size relationship between the three parameters given above is not absolute, and the equation will change accordingly as the changing of the relative size of the parameters and we just present the numerical results later in the Sect. 3. Apparently, the arrival rate of the above will be the same with the bottleneck model proposing by Zhang when the bus mode is absent.

2.2 Bus Mode

2.2.1 Model Introduction

Then, we get to discuss the bus mode. The model for bus mode is the extension work proposed by Huang et al. [4]. The total cost of a commuter who chose the ith bus run is

$$\begin{aligned} C_2(i) = \alpha\left[T_2 + T_{w\_b}(t_i)\right] + \delta_b(t_i) \\ + \left[T_2 + T_{w\_b}(t_i)\right] cong(n_i) + P_2 \end{aligned} \tag{17}$$

To make the problem easier to describe, when $t_i \notin \left[t_z^e, t_z^l\right]$, the waiting time and schedule delay cost of bus mode are same with the auto mode. However, when $t_i \in \left[t_z^e, t_z^l\right]$, it's assumed that the bus runs can go through the bottleneck quickly, so the waiting time in bottleneck is lower. Then we get

$$T_{w\_b}(t_i) = \begin{cases} T_{w\_c}(t_i), & t_i \notin \left[t_z^e, t_z^l\right] \\ T_{w\_b}(t_i), & t_i \in \left[t_z^e, t_z^l\right] \end{cases} \tag{18}$$

And the schedule delay costs of the bus mode,

$$\delta_b(t_i) = \begin{cases} \delta_c(t_i), & t_i \notin \left[t_z^e, t_z^l\right] \\ \delta_b(t_i), & t_i \in \left[t_z^e, t_z^l\right] \end{cases} \tag{19}$$

Similarly, the bus runs is labeled as $i \in \Phi = \{a, \ldots, 2, 1, 0, -1, -2, \ldots - b\}$. Where a and b is large so that all commuters can select the bus runs as they want. $i(> 0)$ denotes that the run arrive at the W early, $i(< 0)$ means late, and $i(= 0)$ is on time. Following the rule FIFO, there is only one bus run arrive at (W) on time. The

arrival time at B for the i th bus run is

$$t_i = \hat{t}_b - \frac{i}{f}, i \in \Phi \tag{20}$$

Every bus commuter will always minimize his/her own travel cost by chose a bus run which depart earlier or later to avoid a high body congestion and long waiting time. It is assumed that there are ω ($\omega \leq a + b + 1$) batches bus run which transport at least one commuter in the equilibrium. Each bus run that takes commuters bears the identical and lowest travel cost.

In conclusion, the equilibrium travel cost can be expressed as

$$\begin{cases} C_2(i) = C_2, \, if \; n_i > 0 \\ C_2(i) \geq C_2, \; \text{otherwise} \end{cases}, i \in \Phi \tag{21}$$

Here C_2 is the total travel cost of the each bus mode commuter in the equilibrium. For given P_2, f and the N_1 auto mode commuters, it is easy to figure out N_2 is the optimum solution of the problem and the optimal solution is unique and can be figured out.

2.2.2 Algorithm for the Equilibrium of the Bus Mode

Huang proved that all bus runs which arrive at B during the rush hour have positive passenger flows and we won't going to prove that here. However, the bus runs which arrive early or late may not be empty too, and we need to confirm the numbers of the passengers on these bus runs. Combining the Eqs. (17)–(19), we can get the equation as

$$C_2 - P_2 - \alpha T_2 - \left[T_2 + T_{w\_b}(t_i)\right] cong(n_i) = \begin{cases} \beta(\beta\gamma\xi/\alpha + i/f), & t_i < t_b^e \\ \beta\gamma\xi, & t_i \in \left[t_b^e, t_z^e\right] \\ \beta(\beta\gamma\xi/\alpha + i/f), & t_i \in \left[t_z^e, \hat{t}_b\right] \\ -\gamma(\beta\gamma\xi/\alpha + i/f), & t_i \in \left[\hat{t}_b, t_z^l\right] \\ \beta\gamma\xi, & t_i \in \left[t_z^l, t_b^l\right] \\ -\gamma(\beta\gamma\xi/\alpha + i/f), & t_i > t_b^l \end{cases} \tag{22}$$

Let $b_{ch\_f}, -b_{ch\_l}$ the first and last bus run chosen by commuters. As the result of that the bus runs are the critical value, one can get that $cong\left(n\left(b_{ch\_f}\right)\right) = cong\left(n\left(b_{ch\_l}\right)\right) = 0$, then we have,

$$\begin{cases} b_{ch\_f} = \left[\left(\frac{C_2}{\beta} - \frac{\beta\gamma\xi}{\alpha}\right) f\right]^c - 1, \\ b_{ch\_l} = \left[\left(\frac{C_2}{\gamma} + \frac{\beta\gamma\xi}{\alpha}\right) f\right]^c - 1, \end{cases} \tag{23}$$

where $[x]^c$ is the smallest integer that not less than x. Therefore, with the definition in the Sect. 2.2.1, we can derive that $\omega = b_{ch\_f} + b_{ch\_l} + 1$.

Next let b_{bo_f} and $-b_{bo\_l}$ be the first and last bus runs which arrive at B in the rush hour except the time of bus lane ($\left[t_b^e, t_z^e\right] \cup \left[t_z^l, t_b^l\right]$), we can easy to have

$$\begin{cases} b_{bo\_f} = \left[\left(\hat{t}_b - t_b^e\right) f\right]^f = \left[\frac{(\alpha-\beta)\gamma\xi}{\alpha} f\right]^f \\ b_{bo\_l} = \left[\left(t_b^l - \hat{t}_b\right) f\right]^f = \left[\frac{(\alpha+\gamma)\beta\xi}{\alpha} f\right]^f \end{cases} \tag{24}$$

where $[x]^f$ is the largest integer that not larger than x.

Similarly, Let b_{zy_f} and $-b_{zy\_l}$ be the first and last bus runs which arrive at B in the bus lane time ($\left[t_z^e, t_z^l\right]$), we have

$$\begin{cases} b_{zy\_f} = \left[\left(\hat{t}_b - t_z^e\right) f\right]^f \\ b_{zy\_l} = \left[\left(t_z^l - \hat{t}_b\right) f\right]^f \end{cases} \tag{25}$$

Employing Eq. (22) and the necessary parameters given by Eqs. (23)–(25), one can obtain the numbers of the passengers corresponding to the each bus run, as the Formula (26).

$$n_i = \begin{cases} 0, & b_{ch\_f} < i \le a \\ cong^-\left(\frac{C_2 - P_2 - \alpha T_2 - \beta(\beta\gamma\xi/\alpha + i/f)}{T_2}\right), & b_{bo\_f} < i \le b_{ch_f} \\ cong^-\left(\frac{C_2 - P_2 - \alpha T_2 - \beta\gamma\xi}{T_2 + T_{w\_b}(t_i)}\right), & b_{zy\_f} \le i \le b_{bo\_f} \\ cong^-\left(\frac{C_2 - P_2 - \alpha T_2 - \beta(\beta\gamma\xi/\alpha + i/f)}{T_2}\right), & 0 \le i \le b_{zy\_f} \\ cong^-\left(\frac{C_2 - P_2 - \alpha T_2 + \gamma(\beta\gamma\xi/\alpha + i/f)}{T_2}\right), & -b_{zy\_l} \le i \le 0 \\ cong^-\left(\frac{C_2 - P_2 - \alpha T_2 - \beta\gamma\xi}{T_2 + T_{w\_b}(t_i)}\right), & -b_{bo\_l} \le i \le -b_{zy\_l} \\ cong^-\left(\frac{C_2 - P_2 - \alpha T_2 + \gamma(\beta\gamma\xi/\alpha + i/f)}{T_2}\right), & -b_{ch\_l} \le i \le -b_{bo\_l} \\ 0, & -b \le i \le -b_{ch\_l} \end{cases} \tag{26}$$

Obviously, employing the Eq. (23) into conservation condition and one can obtain that:

$$\sum_{-b_{ch\_l} \le i \le ch\_f} n_i = N_2 \tag{27}$$

It is easy to found that the equation and some parameters such as ξ, b_{ch_f}, b_{ch_l} are relate to the value of N_1 and N_2, and the C_2 can be written as an unknown function of N_1 and N_2 when all other parameters are given. Once $C_2(C_1 = C_2)$ is determined, all the numbers of passengers on each bus runs can be figure out via the equations above.

2.3 Energy Consumption Model

For the energy consumption model, it's special because of the discretization of the numerical simulation, and the instantaneous velocity of the vehicles (major decision variable considering energy consumption) can't be obtained. Therefore, we take the travel time (free flow time and waiting time) as decision variable to measure energy consumption in this paper.

From home to the bottleneck, vehicles enjoy a state of free flow (more than 60 km/h) and low energy consumption. In addition, the bus runs without a queue when setting the bus lane will enjoy free flow in whole commute and the carrying rate of the bus have little effect on energy consumption [7, 8]. Therefore, in this paper, it's assumed that the energy consumption for the private cars per unit free flow time $EC_{c-f} = 0.08$ (Con/min), and for bus runs $EC_{b-f}(n_i) = 0.24$ (Con/min).

During the queuing process, the vehicles are walking and stopping all the time, and the velocity between vehicles are basically equal (less than 40 km/h). The vehicles have to endure a high energy consumption when the velocity, and the carrying rate of the bus runs have a large impact on that. It's assumed that the energy consumption for private cars per unit waiting time $EC_{c-w} = 0.13$ (Con/min). For the bus runs, the energy consumption per unit waiting time be calculated as Eq. (28) by polynomial fitting ($K = 80$ is the maximal number of passenger that one bus run can take),

$$EC_{b-w}(n_i) = -0.064\left(\frac{n_i}{K}\right)^3 + 0.112\left(\frac{n_i}{K}\right)^2 + 0.04\left(\frac{n_i}{K}\right) + 0.24 \tag{28}$$

With the assumption above, then we have the total energy consumption as

$$\begin{aligned} EC = {} & N_1\left(EC_{c-f}\right) + N_2\left(EC_{b-f}(n_i)\right) \\ & + \sum\left(T_{w\_c}(t)(EC_{c-w})\right) + \sum_{i\in\Phi}\left(T_{w\_b}(t_i)EC_{b-w}(n_i)\right) \end{aligned} \tag{29}$$

3 Numerical Results

In this section, we will present some numerical examples to demonstrate the model. In order to insure the credible of the results, we adapt some parameters form the paper of Huang. The parameters of the numerical example are listed in the Table 2. Similarly, the unit of 2 min is also used for seemingly intuitively.

Likewise, we adopt the following in-carriage congestion function:

$$\text{cong}(n_i) = -5\ln\left(1 - (n_i/k)^{(1/4)}\right) \tag{29}$$

Table 2 The parameters used in the simulation

Variable	α, β, γ	N	S, S_b	λ	f	t_w	μ	$[t_z^e, t_z^l]$	T_1, T_2	P_1, P_2
value	20,15,30	10,000	2000, $s/3$	3	30	8:00	0.05	[7:30,8:30]	0.5,0.7	20,2
unit	HK\$/h	coms	CEQs/h	CEQs/bus	bus/h	AM		AM	h	HK\$

3.1 Results with Peak-Only Bus Lane and Without Bus Lane

Figures 1 and 2 depict the arrival rate at B and W without and with bus lane respectively. The main results from Fig. 1 are $\left[t_c^e, t_c^l\right] = \left[t_b^e, t_b^l\right] =$ [6:20, 8:50], $\hat{t}_c = \hat{t}_b =$ 6:46, $b_{ch\_f} = 38$, $b_{ch\_l} = 74$, $b_{bo\_f} = 13$, $b_{bo\_l} = 62$, $N_1 = 4818$, $N_2 = 5182$, $C_1(C_2) = 55.67\left(HK\$\right)$, $EC = 39229$(Cons).

Figure 1a show that the arrival rate of auto mode at B is 264 (cars/2 min) in [6:20, 6:46], 24 in [6:46, 8:50] and zero for other time. The arrival rate for bus mode represents the number of the passenger on each bus run. There is a local minimum for the bus mode at $\hat{t}_b =$ 6:46 (lower in-carriage cost to avoid high travel cost). Figure 1b shows the bus runs have to wait a long time at W due to the high arrival rates at B during [6:20, 6:46] and this will cause a large energy consumption.

The main results in Fig. 2 are $\left[t_c^e, t_c^l\right] =$ [6:34, 8:44], $\left[t_b^e, t_b^l\right] =$ [6:34, 7:30] $\cup$ [8:30, 8:44], $\hat{t}_c =$ 6:56, $\hat{t}_b =$ 8:00, $t_z^e =$ 6:48, $t_z^l =$ 8:10, $b_{ch\_f} = 71$, $b_{ch\_l} = 35$, $b_{bo\_f} = 43$, $b_{bo\_l} = 21$, $b_{zy\_f} = 15$, $b_{zy\_l} = 15$, $N_1 = 3603$, $N_2 = 6397$,

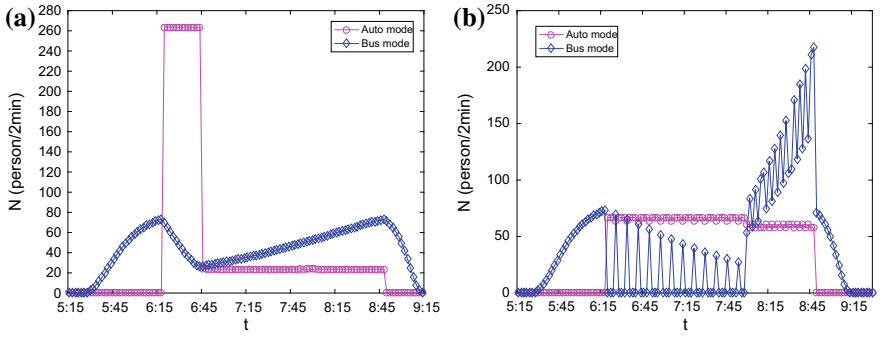


Fig. 1 Arrival patterns without bus lane **a** at B and **b** at W. ($N = 10{,}000$, $S_b = s/3$, $P_1 = 20$, $P_2 = 2$, $f = 30$)

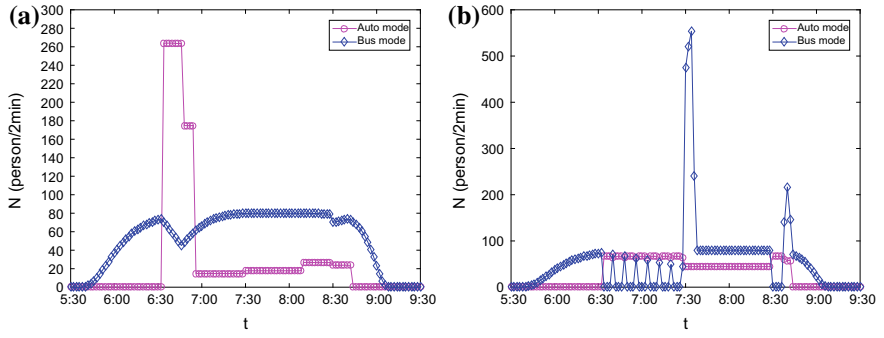


Fig. 2 Arrival patterns with bus lane **a** at B and **b** at W. $\left(\left[t_z^e, t_z^l\right] = [7{:}30, 8{:}30]\right)$

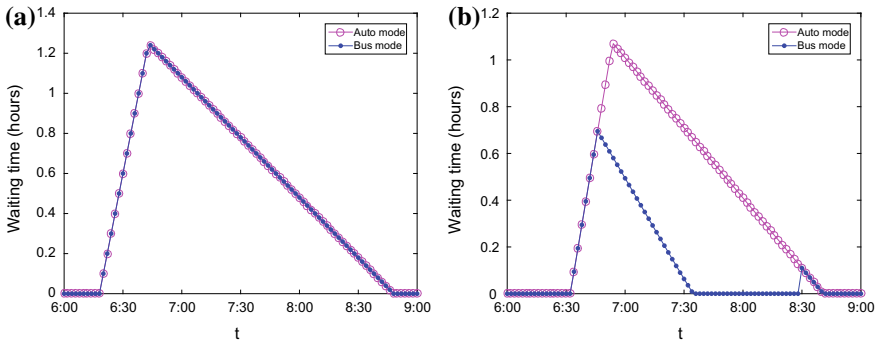


Fig. 3 Waiting time in the queue **a** without bus lane and **b** with bus lane. $\left(\left[t_z^e, t_z^l\right] = [7{:}30, 8{:}30]\right)$

$C_1 = C_2 = 52.37\text{(Hk\$)}, CE = 24883\,\text{(Cons)}$. Then one can get the relationship $t_c^e < \hat{t}_z^e < \hat{t}_c < t_z^e < \hat{t}_z^l < t_z^l < t_c^l$ which is corresponds to Case (i) given in the Sect. 2.1. Obviously, the setting of the bus lane will attract many commuters to choose the bus mode and the total costs and energy consumption decrease.

Figure 2a shows the arrival rate of the auto mode at B is 264 (cars/2 min) in [6:34, 6:48], 175 in [6:48, 6:56], 14 in [6:56, 7:30], 18 in [7:30, 8:10], 27 in [8:10, 8:30], 24 in [8:30, 8:44] and zero for other time. The results are consistent with Case (i) when substituting the variable above. For the bus mode, Fig. 2a shows the bus runs are almost full of passengers because of the privilege of using the bus lane. In addition, there are two local minimums at $\hat{t}_c = 6{:}56$ (longest waiting time) and $t_z^l = 8{:}30$ (join the queue again). Figure 2b verifies the bus runs will first wait and then pass the bottleneck quickly when the bus lane is available and endure waiting time again when unavailable.

As Fig. 3a shows, the waiting time of the two modes are definitely different when the bus lane is set or not. In the presence of the bus lane, there are more commuters prefer to choose the bus mode who have a low energy consumption rate and privilege to pass the bottleneck quickly to travel, and thus decrease the total energy consumption to some degree.

3.2 The Case with Different Constraints

3.2.1 The Results Corresponding to the Case (ii)

In this section, we conduct the simulations after we decrease the length of the peak-only bus lane open period to $\left[t_z^e, t_z^l\right] = [7{:}30, 8{:}10]$.

The results from Fig. 4 are $\left[t_c^e, t_c^l\right] = [6{:}34, 8{:}44]$, $\left[t_b^e, t_b^l\right] = [6{:}34, 7{:}30] \cup [8{:}10, 8{:}44]$, $\hat{t}_c = 6{:}56$, $\hat{t}_b = 8{:}00$, $\hat{t}_z^e = 6{:}48$, $\hat{t}_z^l = 7{:}20$, $b_{ch\_f} = 70$, $b_{ch\_l} = 35$, $b_{bo\_f} = 43$, $b_{bo\_l} = 22$, $b_{zy\_f} = 15$, $b_{zy\_l} = 5$, $N_1 = 3821$, $N_2 = 6179$, $C_1 =$

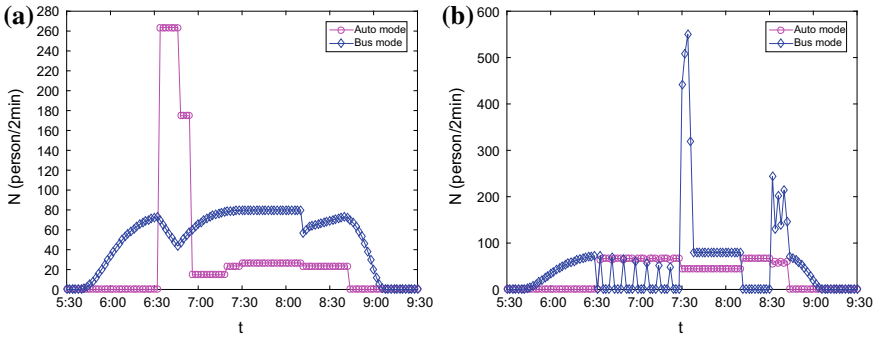


Fig. 4 Arrival patterns of two modes **a** at B and **b** at W. ($N = 10{,}000$, $S_b = s/3$, $P_1 = 20$, $P_2 = 2$, $f = 30$, $\left[t_z^e, t_z^l\right] = [7{:}30, 8{:}10]$)

$C_2 = 52.20(\text{HK\$})$, $CE = 25964$ (Cons). Notice that the total cost and commuters who choose auto mode increase as the open period of the bus lane decrease. However, the total energy consumption increase because more private on the road and thus cause longer waiting time.

Similarly, we can obtain the relationship that $t_c^e < \hat{t}_z^e < \hat{t}_c < \hat{t}_z^l < t_z^e < t_z^l < t_c^l$ which is corresponds to Case (ii) given in Sect. 2.1. Moreover, Fig. 4a shows the arrival rate of auto mode is 264 (car/2 min) in [6:36, 6:50], 174 in [6:50, 6:56], 15 in [6:56, 7:20], 24 in [7:20, 7:30], 27 in [7:30, 8:30], 24 in [8:30, 8:46], and zero for the other time, and these results are consistent with these in the Case (ii). In addition, the local minimum at t_z^l in this is more obvious than the case above because of the increasing of the commuters choosing auto mode, and Fig. 4b there are more bus runs arrive at W during the later stage. Figure 5 shows that there are more bus runs have to endure the waiting time when the bus lane is unavailable which means a lager energy consumption.

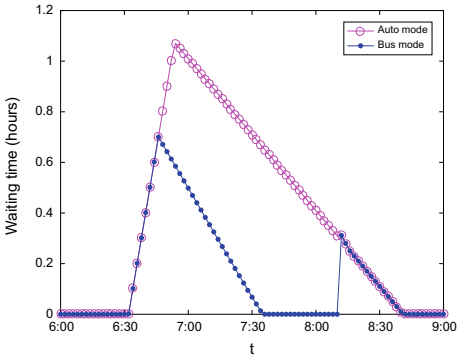


Fig. 5 Waiting time in the queue ($N = 10{,}000$, $S_b = S/3$, $P_1 = 20$, $P_2 = 2$, $f = 30$, $\left[t_z^e, t_z^l\right] = [7{:}30, 8{:}10]$)

3.2.2 The Results Corresponding to the Case (III)

Next, we conduct the simulations with changing several parameters such as $S_b = s/5$, $P_1 = 25$, $\left[t_z^e, t_z^l\right] = [7{:}00, 8{:}30]$ and remain others unchanged.

The main results from Fig. 6 are $\left[t_c^e, t_c^l\right] = [6{:}48, 8{:}36]$, $\left[t_b^e, t_b^l\right] = [6{:}48, 7{:}00] \cup [8{:}30, 8{:}36]$, $\hat{t}_c = 7{:}06$, $\hat{t}_b = 8{:}00$, $\hat{t}_z^e = 6{:}52$, $\hat{t}_z^l = 8{:}22$, $b_{ch\_f} = 73$, $b_{ch\_l} = 36$, $b_{bo\_f} = 36$, $b_{bo\_l} = 18$, $b_{zy\_f} = 30$, $b_{zy\_l} = 15$, $N_1 = 2965$, $N_2 = 7035$, $C_1(C_2) = 53.66(\text{HK\$})$, $CE = 19177(\text{Cons})$. Then we can get the relationship of the variable is corresponds to the Case (iii). Equally, the arrival rete of the auto mode are consistent with the results in Case (iii) and we will not go in detail here. However, Fig. 6 shows that the bus runs are almost full of passengers during the peak period, and there are two local minimum which are not obvious. There are few bus runs need to wait in the queue and Fig. 7 also shows this characteristic and this verifies that privilege for the bus runs by the bus lane.

In summary, we can verify that the commuters will prefer to choose the bus mode to travel from home to the workplace which can reduce their total travel cost and the

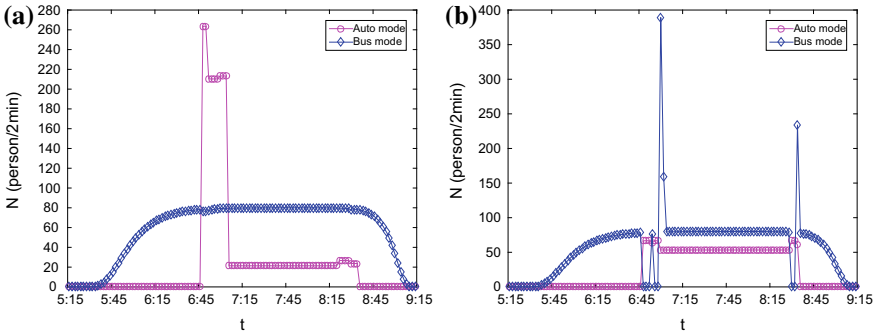


Fig. 6 Arrival patterns of two modes **a** at B and **b** at W. ($N = 10{,}000$, $S_b = S/5$, $P_1 = 25$, $P_2 = 2$, $f = 30$, $\left[t_z^e, t_z^l\right] = [7{:}00, 8{:}30]$)

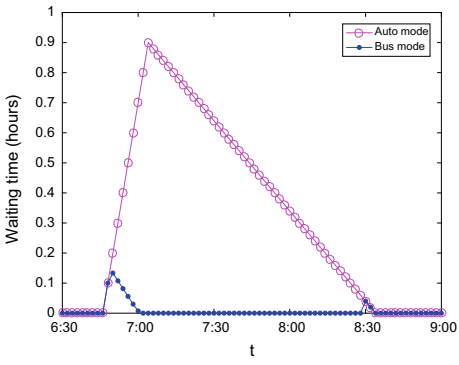


Fig. 7 Waiting time in the queue ($N = 10{,}000$, $S_b = S/5$, $P_1 = 25$, $P_2 = 2$, $f = 30$, $\left[t_z^e, t_z^l\right] = [7{:}00, 8{:}30]$)

energy consumption when the bus lanes are activated. In order to better improve the efficiency and environmental protection of the public transportation, next we will conduct some sensitivity analyses of the model parameters.

3.3 Sensitivity Analyses of the Parameters of the Bus Mode

In this section, we will conduct some sensitivity analyses of the variables such as the departure interval, bus lane's fixed dedicated space and the bus lane open period.

Firstly, we analysis the influence by the change of the duration of the bus lane's availability. We can find that the number of commuters who choose the auto mode and the energy consumption will decrease when the bus lane is present for longer time from the Table 3. However, the cost of the commuters will first decrease and then increase as the length of the duration of the bus lane increases. That is to say, the open period should be set properly.

Then we conduct the analyses of the impact by changing the bus lane dedicated space. Table 4 shows the number of auto mode commuters and total energy consumption increase while the cost of commuters decrease as the dedicated space decreases,

Table 3 Sensitivity analyses of bus lane open period with $N = 10{,}000$, $f = 20$, $C_b = S/3$

Bus lane open period (AM)	N_1(commuters)	N_2 (commuters)	$C_1(C_2)$(HK$)	EC(Cons)
8:00–8:00 (*Nobuslane*)	4818	5182	55.67	39,229
7:50–8:10	4061	5939	52.42	28,770
7:40–8:20	3855	6145	52.30	26,562
7:30–8:30	3603	6397	52.37	24,883
7:20–8:40	3470	6530	52.65	24,220
7:10–8:50	3375	6625	52.67	23,197
7:00–9:00	3290	6710	52.70	22,804

Table 4 Sensitivity analyses of the road resource of bus lane with $N = 10{,}000$, $\left[t_z^e, t_z^l\right] =$ [7:30, 8:30], $f = 30$

S_b	N_1	N_2	$C_1(C_2)$	EC
S/2	3500	6500	53.16	23,680
S/3	3603	6397	52.37	24,883
S/4	3739	6261	51.94	25,648
S/5	3783	6217	51.67	25,983
S/6	3810	6190	51.63	26,420
S/7	3848	6152	51.50	26,562

Table 5 Sensitivity analyses of bus dispatch frequency with $N = 10{,}000$, $\left[t_z^e, t_z^l\right] = [7{:}30, 8{:}30]$

f	$S_b = S/3$				$S_b = S/7$			
	N_1	N_2	$C_1(C_2)$	EC	N_1	N_2	$C_1(C_2)$	EC
20	4988	5012	59.49	40,536	5174	4826	58.11	41,506
30	3603	6397	52.37	24,883	3848	6152	51.50	26,562
40	2728	7272	47.52	16,701	2964	7036	46.76	18,059
50	2089	7911	44.00	11,854	2342	7658	43.48	13,192
60	1627	8373	41.34	8863	1886	8114	40.96	10,058
70	1283	8717	39.29	6897	1550	8450	39.13	8106
80	992	9008	37.71	5591	1225	8775	37.34	6425
90	812	9188	36.27	4721	1010	9090	36.06	5381
100	668	9332	35.17	4165	1122	8878	36.74	6369

which means that the bus lane will be crowded when the bus lane dedicated space is low and the commuters prefer to travel by the private cars.

Finally, we investigate the impact of dispatch frequency. Table 5 shows that the commuters who prefer the auto mode, travel costs and total energy consumption dramatically decrease as the dispatch frequency increase. However, as we can see in Table 5, the results turn to increase when the frequency is too large while the bus lane dedicated space is not big enough where the bus lane will be fully load by bus runs and cause the congestion.

4 Conclusions

This paper examines the travel mode and the departure time in a bi-modal system based on the bottleneck-constrained highway with a peak-only bus lane.

In this study, we compare the commute pattern, travel cost and energy consumption when the bus lane is present or not. The main results we find can be generalize as follows: (i). The join of the bus lane will descend the number of commuters who choose the auto mode, the travel cost and energy consumption will decrease in certain degree. (ii). The duration of the peak-only bus lane need to be set properly with the dispatch frequency. If the open period is too large while the frequency is not big enough, the bus lane will be almost empty which will decline the efficiency and cause the unnecessary waste of road resources. However, the energy consumption will decrease as the increase of the number of commuters who choose the bus mode due to the lower per capita energy consumption. Therefore, we need to strike a good balance between economy and green. (iii). For a fixed capacity of bus lane, the number of auto mode commuters will decrease firstly and then increase as the dispatch frequency increases. Similarly, the travel cost and the energy consumption

will show the same trend. Therefore, when setting a bus lane, the managers need to choose a proper bus dispatch frequency to make full use of the bus lane.

The results summarized in the paper can help us to better understand that the governor should make best balance between the factors which would lead to the transportation system better alleviate traffic congestion and reduce energy consumption when setting a Peak-only bus lane. However, the conditions set in this paper are limited, which is not accord with the real life well, so the model can be further improved by considering additional heterogeneity among the commuters such as different work time, different value of time and different workplace.

Acknowledgements This work is supported by National Key R&D Program of China (No. 2018YFB1600900), the National Natural Science Foundation of China (Grants No. 71621001, 71771021), and the Fundamental Research Funds for the Central Universities (Grant No. 2019JBM035).

References

1. Vickrey WS (1969) Congestion theory and transport investment. Am Econ Rev 59(2):251–260
2. Ludovic JA et al (2011) Capacity drops at merges: an endogenous model. Transp Res Part B 45(9):1302–1313
3. Guler S et al (2012) Strategies for sharing bottleneck capacity among buses and cars. Transp Res Part B 46(10):1334–1345
4. Ren HL, Xue Y, Long J, Gao ZY (2016) A single-step-toll equilibrium for the bottleneck model with dropped capacity. Transportmetrica B Transp Dyn 4(2):92–110
5. Huang HJ et al (2007) Modal split and commuting pattern on a bottleneck-constrained highway. Transp Res Part E 43(5):578–590
6. Zhang XN et al (2010) Analysis of user equilibrium traffic patterns on bottlenecks with time varying capacities and their applications. Int J Sustain Transp 4(1):56–74
7. Zhang S, Wu Y, Liu H (2014) Real-world fuel consumption and CO_2, emissions of urban public buses in Beijing. Appl Energy 113(6):1645–1655
8. Silva C, Bravo J, Gonçalves G (2008) Bus public transport energy consumption and emissions versus individual transportation. Transp Land Use, Plann, and Air Qual Congr 147–160

Research on Public Transit Priority Level and Travel Cooperation Level

Chengming Zhu, Zhenhua Mou, Changxi Ma and Yugang Wang

Abstract Public transit priority strategy is a key measure for alleviating urban traffic congestion problem, but there are few researches on public transit priority level. The connotation of public transit priority level is analyzed. Public transit priority level includes travel time priority level, travel expense priority level, travel comfort priority level and comprehensive priority level. Public transit travel time priority level is analyzed at different travel distance, different public transit carrying speed and car travel speed. Questionnaire survey based on traveler's acceptance is conducted. The relationship between travel cooperation level and travel time priority level, travel expense priority level, travel comfort priority level is obtained. The results shows that the key of improving travel cooperation level is to moderately increase car travel expense, increase public transit travel time priority as far as possible and increase public transit travel comfort as far as possible.

Keywords Public transit · Priority level · Acceptability · Travel cooperation level

1 Introduction

Urban traffic problem is very prominent, especially traffic congestion problem. Public transit priority strategy is a key measure for alleviating urban traffic congestion problem. There are many researches on public transit priority.

C. Zhu (✉) · Y. Wang
School of Energy Science and Engineering, Henan Polytechnic University, Jiaozuo City 454000, Henan Province, China
e-mail: zhuchengming@hpu.edu.cn

Z. Mou (✉)
Beijing Key Laboratory of Traffic Engineering, Chaoyang District Paradise Park 100, Beijing, China
e-mail: mouzhenhua@163.com

Shandong Jianzhu University, Fengming Road 1000, Licheng District, Jinan, China

C. Ma
School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_45

Eriksson L Perceived attributes of bus and car mediating satisfaction with the work commute, traveler's satisfaction on using car is higher than satisfaction on using public transit [1]. Abdolmatin S explored mobility equity in a society undergoing changes in travel behavior [2]. Steven Farber et al. studied on space–time mismatch between transit service and observed travel patterns in the Wasatch Front, Utah: A social equity perspective [3]. Patrick Miller et al. analyzed the sustainability performance of public transit [4]. Craig Morton et al. studied on customer perceptions of quality of service in public transport. Attitudes regarding quality of bus service vary significantly across passenger groups, with females having a tendency to exhibit relatively negative opinions regarding the quality of the cabin environment with a similar finding observed in the case of passengers who are looking after the home and family [5]. Chakrabarti S studied on how can public transit get people out of their cars? An analysis of transit mode choice for commute trips in Los Angeles. Careful planning can promote discretionary transit use by attracting existing latent demand and by creating new demand in an era of increasing government interest in transit and growing traffic congestion. Broader positive effects on the travelling public and the environment are much greater than what this study can predict [6]. David Verbich discussed public transit fare structure and social vulnerability in Montreal. individuals residing in marginalized neighborhoods are likely to spend more money on transit fares over the course of a month compared to those residing in wealthy neighborhoods [7]. Zhang Z L discussed trip mode structure and social benefits maximization based on public transit priority [8]. Alex Karner assessed public transit service equity using route-level accessibility measures and public data, which found that using a single method to evaluate the fairness has obvious shortage [9]. Mitja Stiglic discussed enhancing urban mobility: Integrating ride-sharing and public transit. Travel-sharing and seamless integration of public transportation can significantly improve urban travel mobility and travel proportion of public transit [10]. Dea van Lierop studied enjoying loyalty: The relationship between service quality, customer satisfaction, and behavioral intentions in public transit. The findings from this study are used to define areas where transit agencies can develop specific strategies in order to benchmark user satisfaction with the aim of growing patronage among the different groups [11]. Yanshuo Sun discussed the implications of the cost of public funds in public transit subsidization and regulation [12]. Zhao B discussed the coordination between travel mode choice and public transit priority policy [13]. Zhuo J analyzed some misunderstandings of public transit priority development strategy, which including why need to implement public transit priority, what is public transit priority and how to implement public transit priority [14]. Zuo Z Y analyzed the proportion of using car transfer to public transit under different public transit traveling speed relative to car traveling speed [15]. Yang L P studied the decision making process of commuting travel mode choice and analyzed the resistance, power and transfer mechanism of using car transfer to public transit [16].

In the above research, the key problem of public transit priority is that the lowest public transit priority level corresponding with expected public transit using proportion is not determined. And this priority changes with the actual traffic conditions, it can be thought of as dynamic public transit priority. The above problem results in that

Table 1 The matrix of absolute acceptance level and relatively acceptance level

Car	Public transit			
	Accept with pleasure	Acceptable	Difficult to accept	Unacceptable
Accept with pleasure	Not acceptable	Not acceptable	Not acceptable	Not acceptable
Acceptable	Not acceptable	Not acceptable	Not acceptable	Not acceptable
Difficult to accept	Acceptable	Acceptable	Not acceptable	Not acceptable
Unacceptable	Acceptable	Acceptable	Not acceptable	Not acceptable

provided public transit service level can't reach the level which can make traveler initiative to choose public transit and the expected aim of public transit priority can't be realized. Traffic congestion also can't be relieved.

Therefore, public transit priority level must be firstly determined before the measure of public transit priority is determined. The key of public transit priority is as far as possible to enhance the comparative advantage of public transit travel service level relative to car travel service level. Absolute acceptance level and comparative acceptance level is as following Table 1.

It can be seen from above tab, the situation which traveler can accept public transit is seldom. When travel service level of each travel is very poor, the traveler will prefer to choose car and won't give preference to public transit.

2 The Connotation of Public Transit Priority Level

Public transit priority level is the priority level of public transit relative to car in order to make traveler willing to choose public transit with different proportion. Public transit priority level is closely related to travel system condition and the main factor is travel speed. Travelers have different travel cooperation level under different public transit priority level.

3 Public Transit Priority Level Analyses

3.1 Travel Time Priority Level

Public transit travel time priority level is the ratio of public transit travel time and car travel time in order to make traveler willing to choose public transit with corresponding proportion under the certain travel costs and travel comfort. It is noted that travel time here is the entire travel time, that is, travel time is time from O to D.

The most important thing is that what ratio of public transit travel time and car travel time is, traveler are willing to initiative choose public transit.

There are four types of traveler's attitude on travel cost, including unacceptable, difficult to accept, acceptable and accept with pleasure. Travel cost is the cost corresponding with expected travel mode structure.

There are three types of traveler's feelings on travel comfort, including too crowded, more crowded and un-crowded. Comfort is the crowding level of public transit passenger in public transit.

The ratio of public transit travel time and car travel time is different with different travel distance. The public transit travel time priority level will be analyzed under different travel distance and different public transit carrying speed and car carrying speed.

The travel time of using public transit includes time T_{OS} from O to public transit station, waiting time T_W, riding public transit time T_R, transfer time T_H and time T_{SD} from public transit station to D. The travel distance of using public transit includes distance L_{OS} from O to public transit station, riding distance L_R and distance L_{SD} from public transit station to D. Total travel distance is L, public transit carrying speed is V_P, walking speed is V_W, car travel speed is V_C.

Then the travel time of using public transit is $T_P = \frac{L_{OS}+L_{SD}}{V_W} + \frac{L_R}{V_P} + T_W + T_H$, the travel time of car is $T_C = \frac{L}{V_C}$.

Public transit travel time priority level is

$$PD_T = \frac{T_{OS} + T_{SD} + \frac{L_R}{V_P} \times 60 + T_W + T_H}{\frac{L}{V_C} \times 60}. \tag{1}$$

It is found that the shorter walk distance, the higher public transit carrying speed, the shorter waiting time and transfer time is, the higher public transit travel time priority level is.

In practice, the change of walking distance and waiting time is not obvious, but the change of transfer time, public transit carrying speed and car travel speed is obvious. If the total time of T_{OS} and T_{SD} is 15 min waiting time is 3 min, transfer time is 8 min (when travel distance is less than 6 km, transfer is not considered.), the walking distance, which includes L_{OS} and L_{SD}, is 1 km the riding distance of using public transit is $L_R = L - 1$ (km), then public transit travel time priority level with different travel distance, public transit carrying speed and car travel speed is as following Tables 2, 3, 4, 5, 6, 7 and 8.

For ground public transit, public transit carrying speed has direct correlation with car travel speed under current urban road network structure. When public transit carrying speed is low, car travel speed is also low, but it must be significantly higher than public transit carrying speed and travel speed. When public transit carrying speed is high, car travel speed is also high, but it must be significantly higher than public transit carrying speed and travel speed.

In the above table, when car travel speed is lower than public transit carrying speed, it is not reasonable. When public transit carrying speed is 15, 20, 30 km/h,

Table 2 Public transit travel time priority level (travel distance is 3 km)

Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level
15	*15*	*2.2*	20	*15*	*2.0*	30	*15*	*1.8*
	20	2.9		*20*	*2.7*		*20*	*2.4*
	25	3.6		25	3.3		*25*	*3.1*
	30	4.3		30	4.0		*30*	*3.7*
	35	5.1		35	4.7		35	4.3
	40	5.8		40	5.3		40	4.9

Table 3 Public transit travel time priority level (travel distance is 6 km)

Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level
15	*15*	*1.6*	20	*15*	*1.4*	30	*15*	*1.2*
	20	2.1		*20*	*1.8*		*20*	*1.6*
	25	2.6		25	2.3		*25*	*1.9*
	30	3.2		30	2.8		*30*	*2.3*
	35	3.7		35	3.2		35	2.7
	40	4.2		40	3.7		40	3.1

Table 4 Public transit travel time priority level (travel distance is 9 km)

Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level
15	*15*	*1.6*	20	*15*	*1.4*	30	*15*	*1.2*
	20	2.1		*20*	*1.9*		*20*	*1.6*
	25	2.7		25	2.3		*25*	*1.9*
	30	3.2		30	2.8		*30*	*2.3*
	35	3.8		35	3.2		35	2.7
	40	4.3		40	3.7		40	3.1

Table 5 Public transit travel time priority level (travel distance is 12 km)

Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level
15	*15*	*1.5*	20	*15*	*1.2*	30	*15*	*1.0*
	20	1.9		*20*	*1.6*		*20*	*1.3*
	25	2.4		25	2.0		*25*	*1.7*
	30	2.9		30	2.5		*30*	*2.0*
	35	3.4		35	2.9		35	2.3
	40	3.9		40	3.3		40	2.7

Table 6 Public transit travel time priority level (travel distance is 15 km)

Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level
15	*15*	*1.4*	20	*15*	*1.1*	30	*15*	*0.9*
	20	1.8		*20*	*1.5*		*20*	*1.2*
	25	2.3		25	1.9		*25*	*1.5*
	30	2.7		30	2.3		*30*	*1.8*
	35	3.2		35	2.6		35	2.1
	40	3.6		40	3.0		40	2.4

Table 7 Public transit travel time priority level (travel distance is 18 km)

Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level
15	*15*	*1.3*	20	*15*	*1.1*	30	*15*	*0.8*
	20	1.7		*20*	*1.4*		*20*	*1.6*
	25	2.2		25	1.8		*25*	*1.9*
	30	2.6		30	2.1		*30*	*2.3*
	35	3.0		35	2.5		35	2.7
	40	3.5		40	2.9		40	3.1

Table 8 Public transit travel time priority level (travel distance is 30 km)

Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level	Public transit carrying speed (km/h)	Car travel speed (km/h)	Time priority level
15	*15*	*1.2*	20	*15*	*0.9*	30	*15*	*0.7*
	20	1.6		*20*	*1.3*		*20*	*0.9*
	25	2.0		25	1.6		*25*	*1.2*
	30	2.4		30	1.9		*30*	*1.4*
	35	2.8		35	2.2		35	1.6
	40	3.2		40	2.5		40	1.9

the possible maximum public transit travel time priority level with different travel distance is as following Table 9.

As can be seen from above relationship, at the same trip distance, the higher the public transit carrying speed is, the lower public transit priority level is (i.e., the bigger the ratio of the public transit travel time and car travel time is). The ratio of public transit travel time and car travel time show a trend of increase when a series of measures are adopted to synchronously improve the speed of public transit and car. From the perspective of comparative advantage, travelers will be more reluctant to use public transit. It also provides a good explanation for to improve public transit carrying speed and car speed with different improvement degree. That is, generally improving social vehicles and public transit speed is not conducive to enhance positive impression of public transit service level for traveler, then the traveler's feeling comparative service level on public transit will become worse and worse.

Table 9 Max PT travel time priority level with different travel distance and public transit carrying speed

Travel distance (km)	Public transit carrying speed (km/h)	Max travel time priority level	Public transit carrying speed (km/h)	Max travel time priority level	Public transit carrying speed (km/h)	Max travel time priority level
3	15	2.9	20	3.3	30	4.3
6		2.1		2.3		2.7
9		2.1		2.3		2.7
12		1.9		2.0		2.3
15		1.8		1.9		2.1
18		1.7		1.8		2.7
30		1.6		1.6		1.6

As can be seen from above relationship, at the same public transit carrying speed, the higher travel distance is, the higher public transit priority level is (i.e., the smaller the ratio of the public transit travel time and car travel time is). From the perspective of travel time, travelers will be more reluctant to use public transit when public transit travel time and car travel time is compared. This also explains why traveler prefers to use car for shorter distance travel.

Since making traveler to initiative use public transit through the ratio change of travel time can't be realized at current road network structure, it is needed to change road network structure, at least letting public transit carrying speed to be faster than car travel speed. The condition of car using amount has little influence on public transit carry speed can be realized at changed road network structure, that is, car using amount has influence on car travel speed only. Whether traveler chooses car or public transit is determined by the comparison of public transit carrying speed and car travel speed with different car using amount.

3.2 Travel Expense Priority Level

Public transit travel expense priority level is the ratio of public transit travel expense and car travel expense in order to make traveler willing to choose public transit with corresponding proportion under the certain travel time and travel comfort.

From theory analysis, the higher public transit travel expense priority level is, that is, the higher the ratio of public transit travel expense and car travel expense, the higher choice proportion of public transit. But the synergic relationship between public transit ticket fare and car travel expense should be good coordination. The key is to determine rational public transit ticket fare.

3.3 Travel Comfort Priority Level

Public transit travel comfort priority level is the ratio of public transit travel comfort and car travel comfort in order to make traveler willing to choose public transit with corresponding proportion under the certain travel time and travel expense. Here comfort refers to the passenger comfort in public transit or car.

Because car travel comfort remains almost unchanged, public transit travel comfort priority level can be divided into very crowded, moderate crowded and not crowded. The public transit travel comfort priority level which traveler can accepted is directly related to public transit travel time, public transit travel time priority level, public transit travel expense and public transit travel expense priority level.

In general, the longer the public transit travel time is, the higher public transit travel comfort priority level accepted by traveler is, that is, the requirement of traveler to public transit travel comfort is higher. Because public transit travel time is related to travel distance and public transit carrying speed, trip distance and public transit

carrying speed commonly influence the traveler's requirements on public transit travel comfort. At the same time, public transit travel time priority level also influence public transit travel comfort priority level which traveler can accept. In general, the higher public transit travel time priority level is, the lower public transit travel comfort priority level which traveler can accept is. The higher public transit travel expense is, the higher public transit travel comfort priority level which traveler can accept is. The higher public transit travel expense priority level is, the lower public transit travel comfort priority level which traveler can accept is.

3.4 Comprehensive Priority Level

Comprehensive priority level of public transit travel refers to public transit priority level which can be accepted by traveler with comprehensive consideration on public transit travel expense, public transit travel time and public transit travel comfort. Comprehensive priority level of public transit travel can be expressed by the ratio of public transit travel comprehensive cost and car travel comprehensive cost.

Travel comprehensive cost includes travel time cost, travel expense cost and travel comfort cost. The key is how to conduct equivalent conversion among different type cost. Specific conversion value can be determined through analyzing relevant investigations on traveler.

4 Public Transit Priority Level and Travel Cooperation Level

Public transit priority level directly affects travel cooperation level. The relationship between travel cooperation level and travel time priority level, travel expense priority level, travel comfort priority level is analyzed as following.

4.1 Travel Cooperation Level Based on Travel Time Priority Level

Travel cooperation level of different travel time priority level will be analyzed under certain travel expense and travel comfort priority level.

Travel expenses mainly consider car travel expenses, travel comfort mainly considers passengers crowded degree of public transit travel.

Questionnaire survey is conducted in Zhengzhou city. It is found that when car travel expense will only permit traveler to use car with 50% of total travel amount, traveler can't accept travel expense. When car travel expense will permit traveler

to use car with 70% of total travel amount, traveler are difficult to accept travel expense. When car travel expense will permit traveler to use car with 100% of total travel amount, traveler can accept travel expense. When car travel expense will permit traveler to use car with more than 100% of total travel amount, traveler will accept travel expense with pleasure. Passenger crowd degree in public transit is divided into very crowded, moderate crowded and not crowded.

The combination condition of travel expense and travel comfort has twelve types, that is A (not acceptable, very crowded); B (not acceptable, moderate crowded); C (not acceptable, not crowded); D (difficult to accept, very crowded); E (difficult to accept, moderate crowded); F (difficult to accept, not crowded); G (accept, very crowded); H (accept, moderate crowded); I (accept, not crowded); J (accept with pleasure, very crowded); K (accept with pleasure, moderate crowded); L (accept with pleasure, not crowded).

Travel time priority level is the ratio of public transit travel time and car travel time. Travel time is entire travel time. The value of ratio is 3, 2.5, 2, 1.8, 1.5, 1.2, 1, 0.9, 0.8, 0.7, 0.6, 0.5.

The relationship between travel cooperation level and travel time priority level is as following Table 10.

From the above analysis, the following rules can be obtained.

(1) At the same public transit travel comfort, the higher traveler's acceptance level on car travel expense is, the higher public transit travel time priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher public transit travel time priority level is; Conversely, the lower public transit travel time priority level is.
(2) At the same traveler's acceptance level on car travel expense, the lower public transit travel comfort is, the higher public transit travel time priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher public transit travel time priority level is; Conversely, the lower public transit travel time priority level is.
(3) At the same traveler's acceptance level on car travel expense and public transit travel comfort, the longer travel distance is, the higher public transit travel time priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher public transit travel time priority level is; Conversely, the lower public transit travel time priority level is.

From the above analysis, the following conclusion can be obtained. The key of improving travel cooperation level is to moderately increase car travel expense, increase public transit travel time priority as far as possible and increase public transit travel comfort as far as possible.

Table 10 The relationship of travel cooperation level and time priority level (travel cost can't be accepted, bus passengers is very crowded)

Time priority level	3	2.5	2	1.8	1.5	1.2	1	0.9	0.8	0.7	0.6	0.5
A: travel cooperation level	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6
B: travel cooperation level	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.7	0.7
C: travel cooperation level	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.7	0.8	0.8
D: travel cooperation level	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4
E: travel cooperation level	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.5	0.5
F: travel cooperation level	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.5	0.5	0.6
G: travel cooperation level	0	0	0	0	0	0	0	0	0.1	0.1	0.1	0.1
H: travel cooperation level	0	0	0	0	0	0	0.1	0.1	0.1	0.2	0.3	0.4
I: travel cooperation level	0	0	0	0	0	0	0.1	0.2	0.3	0.4	0.5	0.5
J: travel cooperation level	0	0	0	0	0	0	0	0	0.1	0.1	0.1	0.1
K: travel cooperation level	0	0	0	0	0	0	0.1	0.1	0.2	0.2	0.3	0.4
L: travel cooperation level	0	0	0	0	0	0	0	0.1	0.2	0.3	0.4	0.5

4.2 Travel Cooperation Level Based on Travel Expense Priority Level

Travel cooperation level of different travel expense priority level will be analyzed under certain travel time and travel comfort priority level.

Travel time mainly consider the priority level of public transit travel time and car travel time, travel comfort mainly considers passengers crowded degree of public transit travel. Traveler acceptance level on travel time priority level is as following: not acceptable, difficult to accept, accept and accept with pleasure. Passenger crowd degree in public transit is divided into very crowded, moderate crowded and not crowded.

The combination condition of travel time and travel comfort has twelve types, that is A (not acceptable, very crowded); B (not acceptable, moderate crowded); C (not acceptable, not crowded); D (difficult to accept, very crowded); E (difficult to accept, moderate crowded); F (difficult to accept, not crowded); G (accept, very crowded); H (accept, moderate crowded); I (accept, not crowded); J (accept with pleasure, very crowded); K (accept with pleasure, moderate crowded); L (accept with pleasure, not crowded).

Travel expense priority level is the ratio of public transit travel expense and car travel expense. The value of ratio is 30, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1.

The relationship between travel cooperation level and travel expense priority level is as following Table 11.

From the above analysis, the following rules can be obtained.

(1) At the same public transit travel comfort, the higher traveler's acceptance level on car travel time is, the higher travel expense priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher travel expense priority level is; Conversely, the lower travel expense priority level is.
(2) At the same traveler's acceptance level on car travel time, the lower public transit travel comfort is, the higher travel expense priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher travel expense priority level is; Conversely, the lower travel expense priority level is.
(3) At the same travel time and public transit travel comfort, the longer travel distance is, the higher travel expense priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher travel expense priority level is; Conversely, the lower travel expense priority level is.

Table 11 The relationship of travel cooperation level and expense priority level (travel time can't be accepted, bus passengers is very crowded)

Expense priority level	30	20	15	10	9	8	7	6	5	4	3	2	1
A: travel cooperation level	0.5	0.5	0.5	0.3	0.3	0.2	0	0	0	0	0	0	0
B: travel cooperation level	0.5	0.5	0.5	0.3	0.3	0.2	0	0	0	0	0	0	0
C: travel cooperation level	0.5	0.5	0.5	0.3	0.3	0.2	0	0	0	0	0	0	0
D: travel cooperation level	0.5	0.5	0.5	0.3	0.3	0.2	0	0	0	0	0	0	0
E: travel cooperation level	0.5	0.5	0.5	0.3	0.3	0.2	0.1	0	0	0	0	0	0
F: travel cooperation level	0.5	0.5	0.5	0.3	0.3	0.2	0.1	0.1	0	0	0	0	0
G: travel cooperation level	0.5	0.5	0.5	0.3	0.3	0.2	0	0	0	0	0	0	0
H: travel cooperation level	0.6	0.6	0.6	0.4	0.3	0.3	0.2	0	0	0	0	0	0
I: travel cooperation level	0.7	0.7	0.7	0.5	0.5	0.4	0.3	0.1	0	0	0	0	0
J: travel cooperation level	0.5	0.5	0.5	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0	0	0
K: travel cooperation level	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.3	0.3
L: travel cooperation level	0.8	0.8	0.8	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.5	0.5

4.3 Travel Cooperation Level Based on Travel Comfort Priority Level

Travel cooperation level of different travel comfort priority level will be analyzed under certain travel time and travel expense priority level.

Travel expenses mainly consider car travel expense, travel time priority mainly considers the ratio public transit travel time and car travel time. Traveler acceptance level on travel expense is as following: not acceptable, difficult to accept, accept and accept with pleasure. Traveler acceptance level on travel time priority level is as following: not acceptable, difficult to accept, accept and accept with pleasure.

The combination condition of travel expense and travel time has sixteen types, that is A (not acceptable, not acceptable); B (not acceptable, difficult to accept); C (not acceptable, accept); D (not acceptable, accept with pleasure); E (difficult to accept, not acceptable); F (difficult to accept, difficult to accept); G (difficult to accept, accept); H (difficult to accept, accept with pleasure); I (accept, not acceptable); J (accept, difficult to accept); K (accept, accept); L (accept, accept with pleasure); M (accept with pleasure, not acceptable); N (accept with pleasure, difficult to accept); O (accept with pleasure, accept); P (accept with pleasure, accept with pleasure);

Travel comfort is passenger crowd degree in public transit. It is divided into very crowded, moderate crowded and not crowded.

The relationship between travel cooperation level and comfort priority level is as following Table 12.

From the above analysis, the following rules can be obtained.

(1) At the same public transit travel time priority level, the higher traveler's acceptance level on car travel expense is, the higher public transit travel comfort priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher public transit travel comfort priority level is; Conversely, the lower public transit travel comfort priority level is.

(2) At the same traveler's acceptance level on car travel expense, the lower public transit travel time priority level is, the higher public transit travel comfort priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher public transit travel comfort priority level is; Conversely, the lower public transit travel comfort priority level is.

(3) At the same traveler's acceptance level on car travel expense and public transit travel time priority level, the longer travel distance is, the higher public transit travel comfort priority level which make traveler to initiative prefer to choose public transit is; or in order to reach same travel cooperation level, the higher travel comfort priority level is; Conversely, the lower travel comfort priority level is.

Table 12 The relationship of travel cooperation level and comfort priority level (travel cost can't be accepted, travel time can't be accepted)

Comfort priority level	Very crowded	Moderate crowded	Not crowded
A: travel cooperation level	0.5	0.5	0.5
B: travel cooperation level	0.5	0.5	0.5
C: travel cooperation level	0.5	0.6	0.7
D: travel cooperation level	0.5	0.6	0.8
E: travel cooperation level	0.3	0.3	0.3
F: travel cooperation level	0.3	0.3	0.3
G: travel cooperation level	0.3	0.3	0.4
H: travel cooperation level	0.3	0.4	0.7
I: travel cooperation level	0	0	0
J: travel cooperation level	0	0	0.1
K: travel cooperation level	0	0.1	0.2
L: travel cooperation level	0	0.2	0.6
M: travel cooperation level	0	0	0
N: travel cooperation level	0	0	0
O: travel cooperation level	0	0	0.1
P: travel cooperation level	0	0.1	0.2

4.4 Passive Cooperation and Initiatively Cooperation

From the above relationship between public transit priority level and travel cooperation level, when the ratio of public transit travel service level and car travel service level can't make traveler to initiatively choose public transit, in order to reach certain travel cooperation level, the only approach is to greatly increase car using expense. At this scene, traveler cooperation level is passive cooperation level. When the ratio of public transit travel service level and car travel service level can make traveler to initiatively choose public transit, traveler cooperation level is initiatively cooperation level. The following is the scene distribution of initiatively cooperation and passively cooperation (Table 13).

Considering the actual situation, travel time priority level less than 0.7 is difficult to achieve, when public travel comfort is very crowded, traveler is impossible to take initiatively in cooperation and car travel expenses not accepted by traveler is also unreasonable, but car travel expenses accepted with pleasure by traveler is obviously not conducive to enhance travel cooperation level. Therefore, the scenarios of reasonable enhancing traveler initiatively cooperation level is as following: travel time priority level is between 0.7 and 1, the corresponding state of travel expense and travel comfort includes (difficult to accept, moderate crowded), (difficult to accept, not crowded), (accept, moderate crowded) and (accept, not crowded). The most harmonious scene of enhancing traveler initiatively cooperation level is (travel expense

Table 13 The scene distribution of initiatively cooperation and passively cooperation (I.C means initiatively cooperation; P.C means passively cooperation)

(Travel expense, travel comfort)	Travel time priority level			
	≥1.5	≥1	≥0.7	≥0.5
(Not acceptable, very crowed)	P.C	P.C	P.C	I.C
(Not acceptable, moderate crowded)	P.C	P.C	P.C	I.C
(Not acceptable, not crowded)	P.C	P.C	I.C	I.C
(Difficult to accept, very crowed)	P.C	P.C	P.C	I.C
(Difficult to accept, moderate crowded)	P.C	P.C	I.C	I.C
(Difficult to accept, not crowed)	P.C	P.C	I.C	I.C
(Accept, very crowed)	P.C	P.C	P.C	I.C
(Accept, moderate crowded)	P.C	P.C	I.C	I.C
(Accept, not crowded)	P.C	P.C	I.C	I.C
(Accept with pleasure, very crowed)	P.C	P.C	P.C	I.C
(Accept with pleasure, moderate crowded)	P.C	P.C	I.C	I.C
(Accept with pleasure, not crowded)	P.C	P.C	I.C	I.C

can be accepted, travel comfort is not crowded), the second is (travel expense can be accepted, travel comfort is moderate crowded).

It can also be seen from above tab, when travel time priority level exceeds 1, travelers are passively cooperation no matter what travel expense priority level and travel comfort priority level, that is, traveler will passively choose public transit. In order to realize traveler initiatively cooperation through travel intervention, travel time priority level must be less than 1.

In conclusion, the necessary public transit priority level through travel intervention is as following: travel time priority level is less than 1, travel expense priority level is difficult to accept or accept, travel comfort priority level is not crowed or moderate crowded. The key is how to realize public transit travel time priority is less than 1, but this aim can't be obviously realized under current urban road network (including rail transit network). Therefore, space structure, function structure and operational organization model of road network must be adjusted and optimized.

5 Conclusion

Public transit priority level includes travel time priority level, travel expense priority level, travel comfort priority level and comprehensive priority level. At the same trip distance, the higher the public transit carrying speed is, the lower public transit priority level is. Generally improving social vehicles and public transit speed is not conducive to enhance positive impression of public transit service level for traveler. At the same public transit travel comfort, the higher traveler's acceptance level on car

travel expense is, the higher public transit travel time priority level which make traveler to initiative prefer to choose public transit is. At the same traveler's acceptance level on car travel expense, the lower public transit travel comfort is, the higher public transit travel time priority level which make traveler to initiative prefer to choose public transit is. At the same traveler's acceptance level on car travel expense and public transit travel comfort, the longer travel distance is, the higher public transit travel time priority level which make traveler to initiative prefer to choose public transit is. At the same public transit travel comfort, the higher traveler's acceptance level on car travel time is, the higher travel expense priority level which make traveler to initiative prefer to choose public transit is. At the same traveler's acceptance level on car travel time, the lower public transit travel comfort is, the higher travel expense priority level which make traveler to initiative prefer to choose public transit is. At the same traveler's acceptance level on car travel time and public transit travel comfort, the longer travel distance is, the higher travel expense priority level which make traveler to initiative prefer to choose public transit is. At the same public transit travel time priority level, the higher traveler's acceptance level on car travel expense is, the higher public transit travel comfort priority level which make traveler to initiative prefer to choose public transit is. At the same traveler's acceptance level on car travel expense, the lower public transit travel time priority level is, the higher public transit travel comfort priority level which make traveler to initiative prefer to choose public transit is. At the same traveler's acceptance level on car travel expense and public transit travel time priority level, the longer travel distance is, the higher public transit travel comfort priority level which make traveler to initiative prefer to choose public transit is. In order to realize traveler initiatively cooperation through travel intervention, travel time priority level must be less than 1. The key of improving travel cooperation level is to moderately increase car travel expense, increase public transit travel time priority as far as possible and increase public transit travel comfort as far as possible.

References

1. Eriksson L, Friman M, Garling T (2013) Perceived attributes of bus and car mediating satisfaction with the work commute. Transp Res Part A 47:87–96
2. Shirmohammadli A, Louen C, Vallée D (2016) Exploring mobility equity in a society undergoing changes in travel behavior: a case study of Aachen, Germany. Transp Policy 46:32–39
3. Farber S, Ritter B, Fu LW (2016) Space–time mismatch between transit service and observed travel patterns in the Wasatch Front, Utah: a social equity perspective. Travel Behav Soc 4:40–48
4. Miller P et al (2016) Analyzing the sustainability performance of public transit. Transp Res Part D 44:177–198
5. Morton C, Caulfield B, Anable J (2016) Customer perceptions of quality of service in public transport: Evidence for bus transit in Scotland. Case Stud Transp Policy 4:199–207
6. Chakrabarti S (2017) How can public transit get people out of their cars? An analysis of transit mode choice for commute trips in Los Angeles. Transp Policy 54:80–89
7. Verbich D, El-Geneidy A (2017) Public transit fare structure and social vulnerability in Montreal, Canada. Transp Res Part A 96:43–53

8. Zhang ZL, Zhao B, Tian QF (2012) Resident trip mode structure and social benefit maximization based on transit priority. J Highw Transp Res Dev 29(8):127–131
9. Karner A (2018) Assessing public transit service equity using route-level accessibility measures and public data. J Transp Geogr 67:24–32
10. Stiglic M et al (2018) Enhancing urban mobility: integrating ride-sharing and public transit. Comput Oper Res 90:12–21
11. Lierop DV, El-Geneidy A (2016) Enjoying loyalty: the relationship between service quality, customer satisfaction, and behavioral intentions in public transit. Res Transp Econ 59:50–59
12. Sun YX et al (2016) Implications of the cost of public funds in public transit subsidization and regulation. Transp Res Part A 91:236–250
13. Zhao B (2011) Synergetic relation between resident trip mode choice and transit priority. Jilin University, Chang Chun
14. Zhuo J (2013) Misunderstandings on the prioritization of public transit development. Int Urban Plann 28(3):51–52
15. Zuo ZY, Yang GC, Shao CF (2012) Modeling modal shift of car travelers to buses based on public transport priority. J Transp Syst Eng Inf Technol 12(1):124–129
16. Yang LP (2014) The model of commute mode shift from private car to public transport. Beijing Jiaotong University, Beijing

An Intelligent Road Waterlogging Sensor for Traffic Safety: Principle and Algorithm

Qin-jian Li, Feng Chen and Huang-qing Guo

Abstract Road waterlogging affects the behavior of driver-vehicle unit, thus road waterlogging perception technique is highly important to traffic safety. In this work, a pressure-guiding waterlogging perception method is introduced, and this sensor is modeled based on the principle of differential pressure to realize the real-time measurement of the road waterlogging level. In order to decrease non-linear measurement error of this waterlogging sensor under complex road environment, this paper proposed an adaptive correction algorithm according to the principle of data fusion. The experimental results show that this proposed method has much higher stability and measurement accuracy than typical measurement methods of the road waterlogging level.

Keywords Road waterlogging · Level measurement · Error compensation · Data fusion

1 Introduction

With the rapid development of economy and urban construction, the problem of traffic congestion has become increasingly serious. In order to ease traffic pressure and ensure smooth travel, many cities have established a large number of overpasses and under tunnels, and there are still low-lying roads in these cities. However, due to the lagging construction of the urban drainage network system and unreasonable drainage system planning, road waterlogging on the above sections caused by rainstorm will not only causes traffic disruption, but also endangers the safety of personal and property, which seriously threatening traffic safety [1]. Therefore, it is necessary to establish road waterlogging monitoring system to perform waterlogging level

Q. Li · F. Chen (✉) · H. Guo
Department of Automation, School of Information Science and Technology, University of Science and Technology of China, Hefei 230027, China
e-mail: chenfeng@ustc.edu.cn

F. Chen
Anhui Loongsec Sci and Tech Co., Ltd, Hefei 230088, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_46

monitoring on the waterlogging frequently-accumulated road sections in these cities to provide security for traffic safety, urban flood prevention and traffic dispersion. The most important part of the road waterlogging monitoring system is the data acquisition terminal, that is, the waterlogging sensor used to detect the depth of road waterlogging [2].

Waterlogging sensors can be divided into image-based sensor, non-contact sensor and contact sensor according to different methods of measurement [3]. Iwahashi et al. [4] proposed a method of layered video coding especially for the use of monitoring a water channel. Kim et al. [5] completed an embedded system for measuring the waterlogging level by camera video. Even though the above two methods for monitoring the waterlogging level through video image have low requirements on the measurement environment, their algorithm complexity is too high, which affects the extraction speed of the waterlogging level data and cannot meet the requirements for real-time detection of road waterlogging level. Mousa et al. [6] proposed an approach to estimate waterlogging level with a dual ultrasonic/passive infrared urban flood sensor system. These two kinds of non-contact waterlogging sensing methods are all evolved from the range finder. The measurement accuracy is high and the measurement range is large. However, the measurement accuracy may be affected by debris in the road waterlogging and cannot be satisfied the requirements for measuring the road waterlogging level under complex road environment. Xu et al. [7] used contact-sensor electronic water gauges to monitor the waterlogging in low-lying areas of urban roads. The measurement accuracy reached 1 cm, but its resolution was only 0.5 cm, which could not meet the requirements for continuous dynamic detection of the road waterlogging level. Xia [8] designed a multi-point waterlogging monitoring and warning system for urban roads which used submerged pressure sensor as road waterlogging sensor. The measurement accuracy also reached 1 cm. It is a commonly used method for real-time detection of waterlogging level under complex road environment. But the submerged pressure sensor is susceptible to blockage and corrosion of sediment and other contaminants in the process of contacting with road waterlogging, and the measurement accuracy is easily affected by environmental factors such as temperature.

To solve the problem that the submerged pressure sensor is easily damaged when measuring the depth of road waterlogging, and the measurement accuracy is easily affected by environmental factors, a pressure-guiding waterlogging level perception method proposed by Zhang et al. [9] was introduced in this paper. Firstly, the measurement principle of the pressure-guiding road waterlogging sensor was introduced, and the precise mathematical output model between the waterlogging level and the related physical quantity was established, and the source of measurement error was analyzed theoretically. Then, a measurement error compensation algorithm based on neural network combined with the principle of data fusion was proposed to solve the problem of output drift of this pressure-guiding road waterlogging sensor, which enhances its anti-jamming capability to environmental factors and improves its measurement accuracy significantly.

2 The Principle of Pressure-Guiding Road Waterlogging Sensor

2.1 Measuring Principle

The structure diagram of pressure-guiding road waterlogging sensor is shown in Fig. 1.

The pressure-guiding road waterlogging sensor mainly consists of a pressure guiding tube, a silicon piezoresistive pressure sensor and various structural connecting bodies. The entire device forms an airtight system in the measurement process and air escape is not allowed. The high pressure chamber of the pressure sensor is connected to the upper end of the pressure guiding tube, and the low pressure chamber is connected to the atmosphere. The pressure guiding tube is installed perpendicular to the road surface, and the bottom is flush with the road surface. When the road waterlogging level rises, a small amount of water will be pushed into the pressure guiding tube to form a small water column. This water column will seal a certain amount of air into the tube. The bottom pressure of water is transmitted to the pressure sensor in the housing through the air in the pressure guiding tube. The pressure sensor converts the differential pressure signal into 0–5 V standard electric signal after the piezoresistive diaphragm senses the pressure difference between the high and low pressure chambers.

The major difference between this pressure-guiding road waterlogging sensor and traditional submerged road waterlogging sensor is that the pressure-guiding road waterlogging sensor uses the principle of air conduction and converts contact type measurement into non-contact type measurement, which avoids the direct contact of pressure sensor with road water in the measurement process, solves the problem of sensor clogging and corrosion.

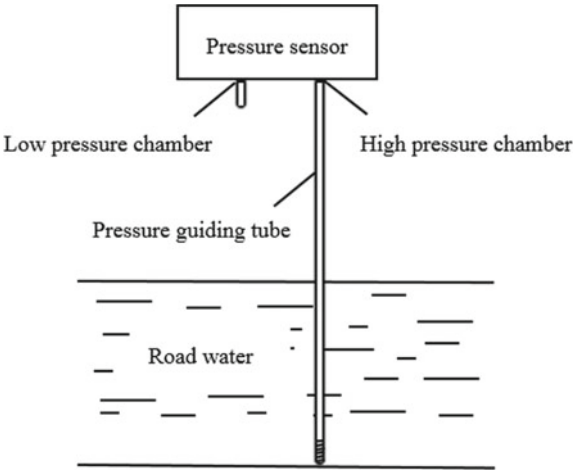


Fig. 1 Structure diagram of pressure-guiding road waterlogging sensor

2.2 Mathematical Output Model

Set the ambient temperature to T (°C), the atmospheric pressure to P_0 (Pa), the waterlogging density to ρ (kg/m^3), the actual depth of road waterlogging is H (m), the height of the small water column in the pressure guiding tube is h (m), the gas pressure in the tube is P (Pa), the amount of air in the tube is n (mol), g is the local gravitational acceleration (m/s^2), and R is the ideal gas constant ($J * mol^{-1} * K^{-1}$), the length of the pressure guiding tube is L (m) and the cross-sectional area is S (m^2).

From the balance of pressure at the bottom of water, there are:

$$P_0 + \rho gH = P + \rho gh \tag{1}$$

From the ideal gas state equation, there are:

$$PS(L-h) = nR(T + 273.15) \tag{2}$$

Substituting Eq. (2) into Eq. (1) and finishing:

$$H = \frac{\Delta P}{\rho g} - \frac{nR(T + 273.15)}{(\Delta P + P_0)S} + L \tag{3}$$

Among them, $\Delta P = P - P_0$, which represents the differential pressure value between the gas pressure in the tube and the atmospheric pressure, calculated by the output voltage of the pressure sensor. Its value changes with the waterlogging level, waterlogging density, ambient temperature and atmospheric pressure in the measurement process.

The Eq. (3) reflects the mathematical function relationship between the actual waterlogging level H and other physical quantity. After obtaining the physical quantities on the right side of the equation by means of approximate values and sensor measurements, the measured value of waterlogging level H' can be calculated by this equation. The difference between H' and H is the measurement error E of this pressure-guiding road waterlogging sensor.

3 The Measurement Error of Pressure-Guiding Road Waterlogging Sensor

3.1 Waterlogging Level Calibration Experiment

In order to analyze and compensate the measurement error of pressure-guiding road waterlogging sensor, the waterlogging level calibration experiment under different ambient temperatures is required. The pressure-guiding road waterlogging sensor used in the experiment uses Loongson 1C as the master chip, Freescale MPX5050 as

Table 1 Waterlogging level calibration experiment data

H/m	*T*/°C						
	10	20	25	30	35	40	50
0	327.78	330.84	332.98	334.71	336.73	338.64	342.8
0.055	387.3	392.88	396.09	397.83	400.03	400.41	403.96
0.101	431.95	436.76	440.24	443.64	447.71	451.66	459.13
0.152	489.37	493.35	496	499.08	502.65	506.14	513.41
0.197	547.21	550.23	552.51	554.94	557.67	560.36	566.33
0.25	610.64	613.77	615.64	617.54	620	622.57	628.05
0.3	666.87	669.89	671.83	674.07	676.47	679.25	684.61
0.35	726.19	729.11	731.12	733.11	735.23	737.76	742.94
0.4	786.28	788.89	790.62	792.33	794.49	796.64	801.75
0.45	840.91	843.72	845.5	847.71	849.81	852.36	857.38
0.5	898.97	901.88	903.4	905.38	907.39	909.81	914.78
0.55	955.81	958.77	960.46	962.42	964.6	966.86	971.99
0.6	1026.353	1028.618	1035.37	1036.938	1038.3	1040.337	1042.35

the pressure sensor and Bosch BMP180 as the temperature sensor. Its operating temperature range is 5–50 °C and measurement range is 0–0.6 m. The pressure guiding tube has a radius of 1.45×10^{-3} m and a length of 1 m. Seven different temperature values are selected within the operating temperature range, and 13 different waterlogging level calibration values are selected within the measurement range. Place the experimental device in the incubator, first adjust the waterlogging level to the first calibration value, and then adjust the temperature of incubator to the first constant value. After the temperature is completely stable, record the corresponding output voltage value of the pressure sensor at this time, and then adjust the temperature to the next value. The above process is repeated until all the depth values of road waterlogging at different selected temperatures have been calibrated. The experimental data obtained are shown in Table 1.

In Table 1, the first line is the ambient temperature T, the first column is the actual depth H of road waterlogging, and the remaining data is the output voltage U (mV) of the pressure sensor.

3.2 Analysis of Measurement Error Sources

From the data in Table 1 in combination with Eq. (3), the measurement error of pressure-guiding road waterlogging sensor in waterlogging level calibration experiments can be calculated. Then, draw the curves of measurement error E with waterlogging level H and ambient temperature T as shown in Fig. 2.

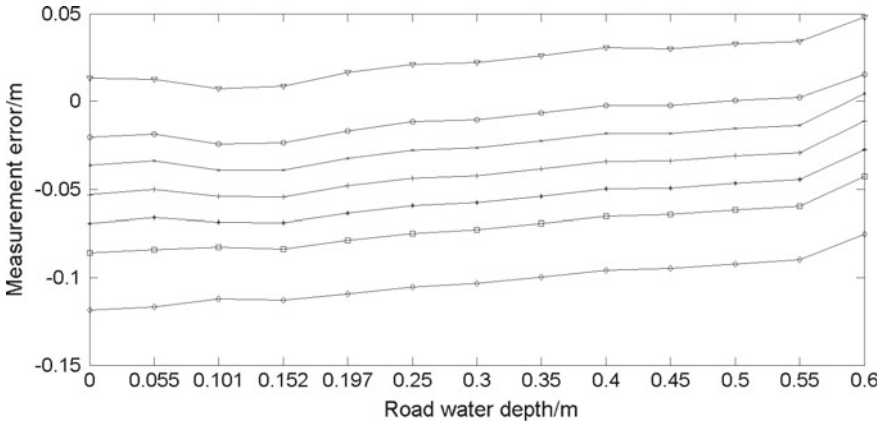


Fig. 2 Curves of measurement error with waterlogging level and ambient temperature

In Fig. 2, the curves are arranged from top to bottom according to the temperature, and are sequentially arranged from 10 to 50 °C.

After analysis, it can be seen that the measurement error of pressure-guiding road waterlogging sensor mainly comes from three aspects. One is the difference between the approximate value and the actual value of some parameters in the mathematical output model. The second is the influence of the change in ambient temperature on the output from the theoretical model in the measurement process. The third is the temperature, zero drift and nonlinear output characteristics of the pressure sensor. Therefore, there is a complex nonlinear relationship between the measurement error and the related physical quantity of pressure-guiding road waterlogging sensor.

4 Measurement Error Compensation for Pressure-Guiding Road Waterlogging Sensor Based on BP Neural Network

4.1 Steps of Algorithm

BP neural network is the most widely used artificial neural network model. It is a unidirectional transmission multi-layer feed-forward network. It trains the network structure through the forward propagation of information and the back propagation of errors, and continuously reduces the distance between the actual output of the network and the expected output [10]. It has the ability to approximate any nonlinear function with arbitrary precision [11]. Therefore, the BP neural network can be used to fit the nonlinear relationship between the measurement error and the relevant physical quantity of pressure-guiding road waterlogging sensor for subsequent measurement error compensation.

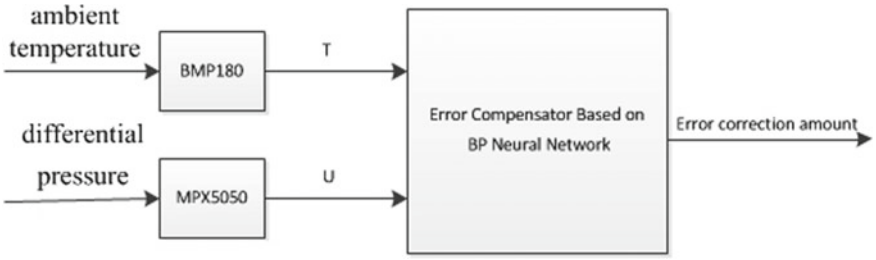


Fig. 3 Data fusion error compensator based on BP neural network

The steps of the measurement error compensation algorithm for pressure-guiding road waterlogging sensor based on BP neural network proposed in this paper are:

(1) Before the measurement, the BP neural network that meets the requirements is constructed by using calibration experimental data. The input parameters of the network during training are the output voltage value U of the pressure sensor in the experimental data and the ambient temperature T measured by the temperature sensor. The output parameters are selected from measurement error E calculated by the experimental data;
(2) During the measurement, the ambient temperature T is obtained by the temperature sensor, and the pressure sensor obtains the output voltage U;
(3) U and T are substituted into mathematical output model to obtain waterlogging level measured value H';
(4) U and T are input into the trained BP neural network, and data fusion of two sensors is performed to obtain an error correction amount E';
(5) The corrected waterlogging level measurement value is obtained by subtracting the error correction amount E' from the measured value H' before compensation.

Data fusion error compensator based on BP neural network aforementioned in steps of algorithm is shown in Fig. 3.

Therefore, the key of this error compensation algorithm is how to adjust the training parameters and the neural network structure to train the BP neural network model that meets the requirements, so that the output measurement error correction amount E' can approximate the true measurement error E.

4.2 Constructing BP Neural Network to Meet the Requirements

In the 91 sets of sample data obtained by calibration experiments, a representative set of 10% (9 sets) of data are selected as a test set (Table 2) for performance testing

Table 2 Selected test set data

Waterlogging level H/m	Ambient temperature T/°C	Output voltage U/mV
0.101	10	431.95
0.152	20	493.35
0.197	25	552.51
0.25	30	617.54
0.3	35	676.47
0.35	40	737.76
0.4	50	801.75
0.45	35	849.81
0.5	30	905.38

of trained neural networks. The remaining 90% (82 sets) data are used as a training set for training neural networks.

In order to speed up the convergence of the training network to achieve a better function fitting effect, the sample data needs to be normalized, and the normalized data are between -1 and $+1$.

The BP neural network structure design in this paper includes one input layer network, several hidden layer networks and one output layer network. The input layer contains 2 neurons and the output layer contains 1 neuron. Considering that it needs to be transplanted into the embedded board when the network training is completed, the hidden layer structure should be as simple as possible. So this paper sets the hidden layer as a single layer. For a single hidden layer BP neural network, the effect of network training is mainly determined by the number of neurons in the hidden layer [12].

To find the suitable number of neurons in the hidden layer, it needs to be confirmed through experiments. In this model, the hidden layer transfer function uses the logsig function, the output layer transfer function uses the pureline function, the training function is the trainlm function, the learning function is the leamgdm function, the learning rate is set to 0.01, the number of fixed iteration is set to 1000, and the target error is set 1×10^{-4} After the training ends, the network can be well fitted for the training samples, but there is a large fitting error in the data that is not in the training set. In other words, the quality of neural network training results depends on its generalization ability. The test set data are substituted into the trained BP neural network for testing, and the performance of the BP neural network on the test set when the number of hidden layer neurons changed is shown in Table 3.

The target of BP neural network training is mean square error. Although it can reflect the effect of training, it is necessary to add the maximum absolute error indicator to reflect the measurement accuracy of pressure-guiding road waterlogging sensor at this time. As can be seen from the data in Table 3, when the number of neurons in the hidden layer increases, the accuracy increases first and then decreases, indicating that the training effect of the neural network has undergone a process from

Table 3 Influence of the number of hidden layer neurons on training effect

Hidden layer neurons	Mean square error	Maximum absolute error
6	1.13E−03	0.0072
10	1.65E−04	0.0021
14	5.53E−05	0.0015
18	2.80E−04	0.0029
22	4.50E−04	0.0033

underfitting to overfitting. In addition, when the neural network is transplanted to the embedded board, the increase of the number of neurons in the hidden layer will increase the complexity of the system and prolong the calculation time. Considering the trade-off between measurement accuracy and computation speed, this paper ultimately determines that the single hidden layer of the BP network contains a total of 10 neurons.

5 Experimental Results and Discussion

Table 4 shows the output in the test set of this trained BP neural network when the hidden layer contains 10 neurons.

The data in Table 4 shows that on the test set: the maximum absolute error of the measurement results before error compensation is 9.64 cm and the maximum relative error is 24.1%; the maximum absolute error of the measurement result after error compensation by this trained BP neural network model is 2.1 mm and the maximum relative error is 1.78%. Both the absolute error and the relative error are reduced by one order of magnitude. This shows that the error compensation algorithm proposed

Table 4 Training effect when the number of hidden neurons is 10

H/m	E before compensation/m	Relative error before compensation/%	E'/m	E after compensation/m	Relative error after compensation/%
0.101	0.0069	6.83	0.0087	−0.0018	1.78
0.152	−0.0242	15.92	−0.0229	−0.0013	0.86
0.197	−0.0331	16.8	−0.0352	0.0021	1.07
0.25	−0.0444	17.76	−0.0456	0.0012	0.48
0.3	−0.0583	19.43	−0.0562	−0.0021	0.7
0.35	−0.0699	19.97	−0.0694	−0.0005	0.14
0.4	−0.0964	24.1	−0.0972	0.0008	0.2
0.45	−0.0498	11.07	−0.0494	−0.0004	0.09
0.5	−0.0316	6.32	−0.0323	0.0007	0.14

in this paper reduces the interference of various error factors, including ambient temperature, to the measurement results of pressure-guiding road waterlogging sensor. After measurement error compensation, the measurement accuracy of this pressure-guiding road waterlogging sensing method proposed in this paper is improved to 3 mm, which is an order of magnitude higher than the measurement accuracy (1 cm) of the typical measurement method using the submerged pressure sensor.

6 Conclusion

This paper introduced a pressure-guiding road waterlogging perception method and established its mathematical output model. Aiming at the problems of non-linear error in the measurement results of pressure-guiding road waterlogging sensor, a measurement error compensation algorithm based on BP neural network combined with the principle of sensor data fusion is presented.

Through the analysis of the experimental results, we can see that this error compensation algorithm effectively enhances the anti-jamming ability of this pressure-guiding road waterlogging sensor to various environmental factors and improves its stability in the measurement process. The measurement accuracy of this pressure-guiding road waterlogging sensor after error compensation is higher than the currently used submerged pressure sensor. Thus, it better satisfies the requirements for the real-time detection of road waterlogging level under complex environment and further guarantees the safety of traffic travel.

Acknowledgements This work was supported by a grant from Anhui Province Foreign Science and Technology Cooperation Project.

References

1. Shi Y (2012) Risk analysis of rainstorm waterlogging on residences in Shanghai based on scenario simulation. Nat Hazards 62(2):677–689
2. Huang Z, Wang S (2013) Based on the water depth sensor network highway measurement system design. Comput Meas Control 21(2):352–354
3. Xu Z, Feng J, Chang H (2017) Research survey of road-water depth measurements. Electron Meas Technol
4. Iwahashi M, Udomsiri S, Imai Y, et al (2007) Water level detection for functionally layered video coding. In: IEEE international conference on image processing. IEEE, pp II-321–II-324
5. Kim J, Han Y, Hahn H (2011) Embedded implementation of image-based water-level measurement system. IET Comput Vision 5(2):125–133
6. Mousa M, Claudel C (2014) Water level estimation in urban ultrasonic/passive infrared flash flood sensor networks using supervised learning. In: IPSN-14 proceedings of the 13th international symposium on information processing in sensor networks. IEEE, pp 277–278
7. Xu H, Liu L (2010) Application of electronic water gauge in monitoring system for urban road waterlogging. Water Resour Informatization 3:015

8. Xia Z (2013) Design of monitoring and alarm system for city stagnant water. Electron Test 5:030
9. Zhang X, Yu S, Zhang J (2016) Lead pressure formula fluviograph:, CN 205561978 U [P]
10. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533
11. Hecht-Nielsen R (1992) Theory of the backpropagation neural network. In: Neural networks for perception. pp 65–93
12. Xia S, Zhao L (2015) Research of multi-sensor information fusion based on BP neural networks. Comput Meas Control

Travel Decision of Shared Bike Based on Subway Transfer

Yongneng Xu, Ren-fei Wu, Qiao Qiao and Zhu-ping Zhou

Abstract Subway is an important part of Public transportation. To increase the attraction of subway, solving the problem of transfer in the beginning and the end of subway is the most important. The paper points out that how shared bikes make it convenient for the transfer of subway by the trip decision of travelers. According to the comparation between shared bikes and other traffic modes, the paper creates a function model by using the difference value between the travel time of two traffic modes. And the paper indicate that search time is an important factor whether most travels will choose shared bikes or not. And then for the improvement of search time, the paper makes a suggestion at both the government and the corporate level, which is aimed to make it convenient to transfer between shared bikes and subway.

Keywords Shared bike · Travel decision · Subway · Transfer

1 Introduction

Recently, with the development of urban transportation, problems such as traffic congestion and environmental pollution have become more serious. Developing public transportation is one of the most effective way to improve the efficiency of urban traffic and reduce environmental pollution. Subway is the backbone of public transport. To encourage more travelers to use subway, solving the problem of the last kilometer in the beginning and the end of subway is most important. Usually the transportation methods related to subway are walking, non-motor vehicle and motor vehicles. The distance influences the travelers' experience of walking. The longer the distance, the worse the experience of travelers. Motor vehicles increase the traffic volume and pollute the environment by exhausting gas. Therefore, motor vehicle is not encouraged. Using non-motor vehicle to link subway is a good choice to achieve the travel of door to door. Especially for those cities that subway is mature, using

Y. Xu · R. Wu (✉) · Q. Qiao · Z. Zhou
Department of Transportation Engineering, Nanjing University of Science and Technology, Nanjing City 210094, Jiangsu Province, China
e-mail: wurenfeidora@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_47

the mode of non-motor vehicle and subway can effectively increase the passenger flow of subway. This mode also makes a large number of travels who originally use motor vehicle choose subway, which can improve traffic efficiency and reduce environmental pollution.

Some innovate enterprises infuse new life into bicycle by using GPS positioning, two-dimensional code scanning, mobile phone mobile payment and so on. According to《the user monitoring report of China's sharing cycle industry in the first quarter of 2017》released by Trustdata (China's mobile Internet data monitoring agency), the number of bicycle users peaking at almost 6 million daily in the first quarter of this year. The rate of chain growth exceeded more than 100%. Daily recharge transactions also soared, and the peak rose up to nearly 100 million. At the same time, the density of shared bicycle shows a rapid growth. The size of first-tier cities have been flat compared with demand. Second and third tier cities still have great potential for development.

Mayer et al. [1] argued that shared bike fills the gap of city's public transportation. It mainly takes the role of connecting with various public transportations. They suggested considering system expansion in the future. Increase the density of residential stations and bikeway, these facilities are an important link to residential area and public transport station. Lin [4] thought that price is the decision-making factor of choosing shared bikes or not. They conducted preference surveys on passengers. They studied the effect of price on travelers' decision with a binary logit model of application and a potential category model on segment specific preferences. Ma et al. [6] also believed that price has effect on travelers' decision. Faghih-imani et al. [2], suggested that consumption of travel time is also an important factor when choosing a vehicle to travel. They compared the traffic flow on the same starting point and destination and explored which is more competitive one, shared bikes or taxis. Godavarthy et al. [3] considered that weather has an impact on using of shared bikes. They carried out an online survey and conducted a regression analysis of these data. Regression analyses were used to estimate the impacts of weather and other factors on bicycle sharing in Fargo.

2 Model of Transportation Mode Selection

2.1 Independent Variable

When analyzing trip characteristic, if the travel time is taken as the abscissa, the result will hide the relationship between choice of travel mode with the travel distance. Therefore, the travel distance is taken as the coordinate.

2.2 Dependent Variable

With the development of economy and public transportation network, more people will choose the mode of public transportation. However, the share rate of public transportation has risen slowly. The problem of the last kilometer is usually solved by walking, private bicycles, private electric bicycles and taxis. The advent of shared bikes injected fresh blood into the means of transportation. At the same time, the shared bikes make travelers to reconsider whether it is a better choice to for the last kilometer. There are many factors to be considered in the choice of traffic mode, mainly including traffic travel cost, travel time consumption, initiative, safety, punctuality, convenience, comfort and accessibility. Among those factors, travel time consumption is an important consideration. So, take travel time consumption as the final decision-making factors of traveler decision-making.

2.3 Comparison Object

Before the advent of shared bikes, travelers usually solve the problem of the last kilometer by walking, using private bicycles or private electric bicycles and taking taxis. After shared bikes appear, the travelers have a new choice. The following models selects the best travel mode according to the minimum travel time. In the travels which are more than 800 m, walking will not be considered because of a serious decline of walking comfort level. Private bicycles and private electric bicycles have limitations and travelers can only take their bicycles in which they park. In the trip chain which appear public transportation, travelers cannot continue to use their private bicycles because they park their bicycle in the beginning of public transportation. For example, a traveler's trip chain is A-B-C-D. He uses his private bike from A to B. The private bicycle is parked at the point B. At the point B, he takes public transportation to the point C. He can only use other traffic mode to point D because his private bike is parked at point B. Therefore, private bicycles will not be considered. So this model takes taxis as a comparison object of shared bicycles.

2.4 Model of Transportation Mode Selection

Ride time of bike share (T_{bike})

$$T_{bike} = t_{quest} + t_{ride} + t_{store} = t_{quest} + \frac{L}{V_{ride}} + t_{store} \tag{1}$$

Ride time of taxi (T_{car})

$$T_{car} = t_{wait} + t_{car} = t_{wait} + \frac{\mathrm{L}}{V_{ride}}(\mathrm{h}) \tag{2}$$

Here:

L The travel distance of rsidents (km)
t_{quest} The time taken by residents to find a shared bikes (h)
t_{ride} The time taken by residents to use a shared bikes (h)
t_{store} The time spent by residents to park and take a shared bikes (h), $t_{store} = 2$ min;
t_{wait} The average time spent by residents waiting for a taxi (h)
t_{car} The time taken by residents to use taxi (h)
V_{ride} The average speed of cycling (km/h)
V_{car} The average speed of a taxi (km/h)

Set $\Delta T = T_{bike} - T_{car}$, Order $\Delta T = 0$, $T_{bike} = T_{car}$
Simultaneous the (1) and (2),

$$\begin{aligned} \Delta T &= b * L - a + t_{quest} \\ a &= t_{wait} - t_{store} \\ b &= \frac{1}{V_{ride}} - \frac{1}{V_{car}} \end{aligned} \tag{3}$$

The structure of travel mode, shown in I of Fig. 1, shows the relationship between travel distance and the share ratio of various traffic modes. When travel distance is shorter, the share ratio of walking is the biggest. With the increase of travel distance, the share ratio of walking decreased gradually, while that of bicycles and public transportation increased gradually. When the travel distance continues to increase,

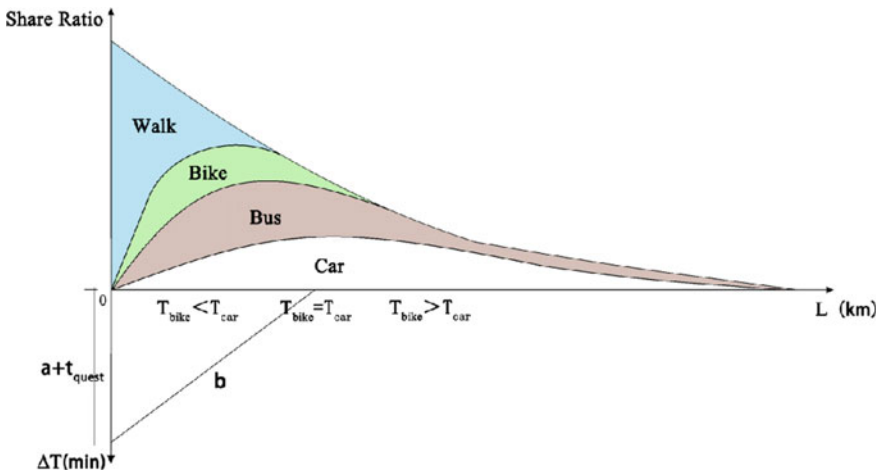


Fig. 1 Structure of residents' travel

the share ratio of walking drops rapidly, and that of bicycles and public transportation decreases slowly, while share ratio of the car increases gradually.

The average speed of various traffic mode in a city is known and the access time of shared bikes under normal circumstances is fixed. If taking the value of them into (3), the values of a and b can be found. Draw the intersection point of $\Delta T = 0$ in the horizontal coordinates of the travel distance in Fig. 1. At the left of the intersection point, $T_{bike} < T_{car}$, travelers tend to choose shared bikes when they travel. At the right of the intersection point, $T_{bike} > T_{car}$ travelers tend to choose taxis when they travel. At the intersection point of $\Delta T = 0$, it is the critical point for travelers to choose shared bikes or taxis when they travel. Explore the influence of search time (t_{quest}), on the consumption of shared bikes through the linear programming of t_{quest}. However, the time taken by travelers to find shared bikes is closely related to the service and technical level of shared bikes. Therefore, the study of t_{quest} is of great significance to the improvement of shared bikes.

3 Case

3.1 Model of Transportation Mode Selection

A total of 1000 questionnaires were sent out in this questionnaire. The cyclists were randomly selected near the subway station in Nanjing by means of all-day delivery. Study the travel characteristics of the bicycle by collecting data.

Through the questionnaire survey, we can get the proportional distribution of the shared bike's travel distance. It is approximately consistent with the travel distribution curve of the bicycle in the traffic trip structure. As shown in Fig. 2, travel distance is taken as abscissa and proportion as ordinate. It can be inferred that the bike share the same characteristic position of the bicycle in the Fig. 1.

Ride time of bike share:

$$T_{bike} = t_{quest} + t_{ride} + t_{store} = t_{quest} + \frac{L}{V_{ride}} + t_{store}(\text{h}) \tag{4}$$

In general, $V_{bike} = 15\,\text{km/h}$, $t_{store} = 2\,\text{min} = 0.033\,\text{h}$

Ride time of taxi:

$$T_{car} = t_{wait} + t_{car} = t_{wait} + \frac{L}{V_{ride}}(\text{h}) \tag{5}$$

In general, $t_{wait} = 15\,\text{min} = 0.25\,\text{h}$, $V_{car} = 40\,\text{km/h}$

Set $\Delta T = T_{bike} - T_{car}$, Order $T = 0$, $T_{bike} = T_{car}$

Simultaneous the (4) and (5),

$$\Delta T = b * L - a + t_{quest}$$

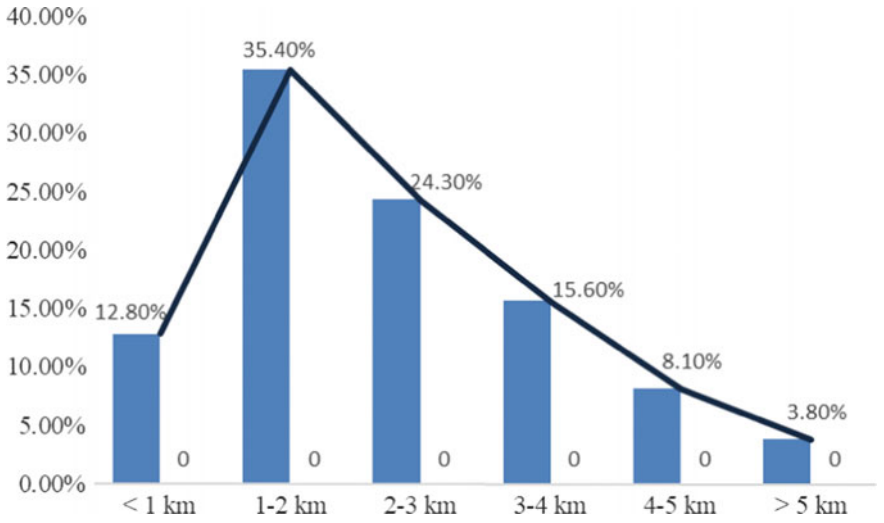


Fig. 2 proportion of different travel distance

$$a = t_{wait} - t_{store} = 0.217$$
$$b = \frac{1}{V_{ride}} - \frac{1}{V_{car}} = 0.041 \tag{6}$$

Set $t_{quest} = 5$ min, $\Delta T = 0$, Order $L = 3.268$
Set $t_{quest} = 5$ min, $L = 0$, Order $\Delta T = 0.134$
Set $t_{quest} = 10$ min, $\Delta T = 0$, Order $L = 1.220$
Set $t_{quest} = 10$ min, $L = 0$, Order $\Delta T = 0.050$

As shown in Fig. 3, When $t_{quest} = 5$ min and $t_{quest} = 10$ min, the linear relation diagram of L and ΔT is shown in the map of resident travel. When t_{quest} increases, the advantages of bike share are getting smaller and smaller. In other words, if want to increase the advantaged distance of bike share and improve its competitiveness, it is necessary to taken by residents to find a shared bike.

In order to reduce search time, it requires bike share companies to put more bikes exactly. However, more bikes available will lead to serious regulatory problems. On this account, vehicle search time cannot decrease without limit. In theory, when t_{quest} is appropriate? According to the questionnaire survey data, most of bike share travel distance is less than 3 km. When ride bike more than 3 km expect for bodybuilding, most of us will feel uncomfortable and want to change other transportation.

So, it's reasonable to study bike travel distance within 3 km.

Bring $L = 3$, $T = 0$ into equation $\Delta T = b * L - a + t_{quest}$

Got $t_{quest} = 0.094$ h $= 5.63$ min

The results shows that the number of t_{quest} is 5.63 min, when the critical point of bike share and taxis is exactly 3 km. In order to increase the competitiveness of bike share, it is necessary to control the time of searching a shared bike within 5.63 min.

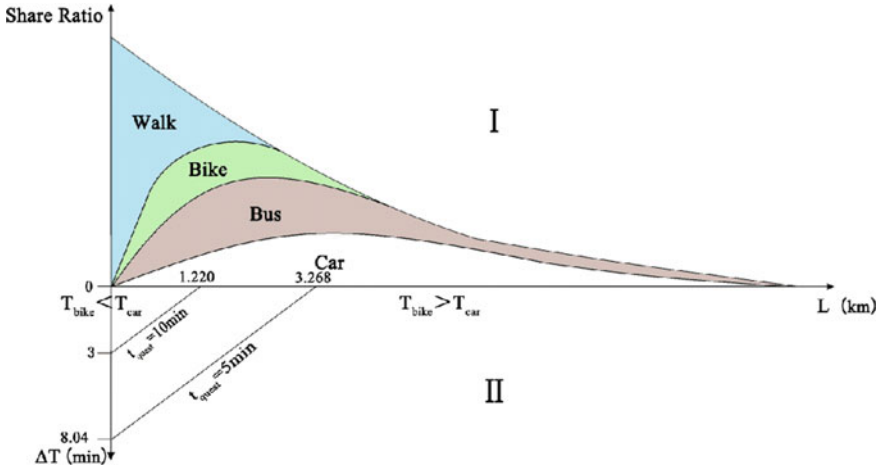


Fig. 3 Relation of distance and time consumption under different t_{quest}

3.2 The Search Time in Reality

As showing in Fig. 4, proportion is taken as abscissa and search time as ordinate. As is shown most residents need to spend 5–15 min looking for a shared bike. After calculation, realistic search time is 10.34 min, longer than theoretical time Thus, the present situation of bike share needs to be improved.

According to the data of questionnaire, it is found that factors affecting search time are not only when and where, but also common problems such as GPS error and

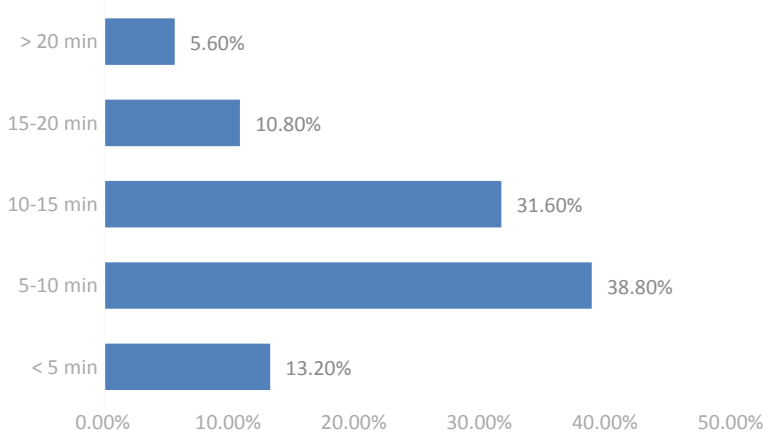


Fig. 4 Specific proportion of search time

vehicle failures. Some Problem caused by excessive investment. Increasing the number of bike share and increasing the density of bike share units is an effective means to reduce vehicle search time. Shorter vehicle search time will lead the advantages of bike share become more prominent. However, the search time cannot be reduced indefinitely, and the excessive amount of data calculated without the data will cause serious waste of resources, which can also lead to a series of problem on parking control.

3.3 Improvement Measures

- Improvement of enterprise
  (a) In-depth research and development on technology is significant. Optimizing vehicle equipment are main approaches to improve the access efficiency of vehicles. Optimize the update efficiency and precision of GPS positioning system also facilitate the user to find the car and reduce the time consumption of the vehicle.
  (b) Improve the operation management system and maintenance immediately in order to minimize the waste of travel time because of vehicle failure.
  (c) Companies of bike share are suggested to open a user active breakdown report function, promoting proper using of shared bike. Through verifying public tip-offs, different degree of punishment should be given according to different severity. Some punishments can be implemented such as deducting credit points, improving the damaged bike use unit price, etc.
  (d) Keep communication with relevant government departments. Through scientific calculation to control the number of bikes on roads. Rather than fighting for market share by increasing the number of shared bike blindly.
- Improvement of government
  (a) Related government organizations should communicate with the companies of bike share regularly and improve the management level of road right and road network. More attention should be paid to adjust the parking area for bikes, according to actual needs. It will help cyclists to park bikes without disturbing other modes of transportation.
  (b) Strengthen the regulation of bicycle traffic and help people have a good habit in riding bike.
  (c) A complete national credit system can be established, unifying the credit behaviors of users in various industries. Malicious damage of shared bike will reduce share through the restriction from credit system.

4 Conclusion

The appearance of shared bike is conducive to subway transfer. In this paper, we build a model to study the decision of using shared bike based on the structure of travelers. As a conclusion, there is a negative correlation relationship between the time of finding shared bike and the advantaged distance of bicycle travel. The smaller the time of finding shared bike, the longer advantaged distance of shared bike compared to the same of taxi. However, due to realistic factors, the search time should be kept at a certain theoretical value when available shared-bike allocation is optimal. Comparing the actual vehicle search time with theoretical search time, the actual search time needed to be ameliorated. According to influence factor of shortcomings, the corrective actions based on the two standpoints of the enterprise and the government are put forward. Thus, the attraction of subway passenger flow transferred from shared bike is improved.

Acknowledgements This research is supported by National Key R&D Program of China under Grant No. 2017YFB1001801 and Fundamental Research Funds for the Central Universities (30917012102). More importantly, we also want to thank the participants in our experiment.

References

1. Mayer M, Brown C (2017) Big blue comes to Jersey City: an analysis of citi bike expansion into New Jersey [C]. Transportation Research Board 96th Annual Meeting, Washington DC. 13
2. Faghih-Imani, A, Anowar, S Miller EJ, Eluru N (2017) Car sharing adoption intention in urban areas: what are the key sociodemographic drivers? Transp Reasahsh Part A, Policy and Practice 101:218–227
3. Godavarthy R, Mattson J, Taleqani AR (2017) Evaluation study of the bike share program in Fargo. North Dakota. p 133
4. Lin JJ, Wang NL, Feng CM (2017) Public bike system pricing and usage in Taipei. Int J Sustain Transp 11(9):633–641
5. Ran L, Li F (2017) An analysis on characteristics and behaviours of travelling by bike-sharing [J]. J Transp Inform Saf 35(6):93–100 (in Chinese)
6. Ma Y, Qin X, Zhang Y et al (2017) Understanding when price can impact the choice of public bicycle: stated performance survey and analysis of Guangzhou public bicycle service [C]. Transportation Research Board 96th Annual Meeting, Washington DC. 19
7. Prieto M, Baltas G, Stan V (2017) Car sharing adoption intention in urban areas: what are the key sociodemographic drivers? Transport Res A: Pol Pract 101:218–227

Evaluation of On-Street Parking Effectiveness Based on Lean Time Management

Jia-li Ge, Wen-hong Lv, Peng-fei Wang, Guo-juan Wang and Lu-li Liang

Research project fund:"Shuangbai" research project in Qingdao (2017-B-24).

Abstract An on-street parking management concept based on lean time management was proposed. The on-street parking evaluation index system based on lean time management is constructed from three aspects: parking fee, parking duration, facility and management. Established an on-street parking time-entropy model based on lean time management, and evaluated the parking effect using the system and model. The main factors influencing the on-street parking effect can be found from the evaluation results, which can provide the basis for further parking planning.

Keywords On-street parking · Lean management · Time-entropy · Evaluation

1 Introduction

Lean management is derived from lean production and summed up from lean production [1]. Lean production is a production organization method proposed by the MIT professors in the international automobile development project (IMVP), which is based on the Japanese Toyota production system (TPS) [2, 3]. In this study, lean time management concept is proposed from time perspective to reduce unnecessary waste of on-street parking time.

On-street parking is usually researched in terms of parking facility supply, parking demand, parking behavior, parking fee, etc. Dai et al. [4] studied the strategy of

J. Ge · W. Lv (✉) · P. Wang · G. Wang · L. Liang
College of Transportation, Shandong University of Science and Technology, Qingdao 266590, China
e-mail: zklwh@sdust.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_48

on-street parking management from the perspective of public security traffic management, and put forward supporting measures, such as rationally set on-street parking spaces; Li et al. [5] proposed urban parking innovation and reform strategies, from the aspects of parking supply-side, parking demand-side, parking supply and demand matching. Alemi et al. [6] used the Poisson distribution model to analyze the parking pricing impact on parking cruising time and distance in San Francisco parking project.

Currently, the major research of traffic evaluation include the socio-economic impact of traffic construction projects and the influence of traffic [7]. For the socio-economic impact evaluation of traffic construction projects, the former Ministry of Construction and the State Planning Commission Investment Research Institute have defined a relatively authoritative evaluation index system in the *Social Evaluation of Transport Project Investment*. The evaluation of traffic influence is a quantitative analysis of the influence of urban land development projects or land-use change when setting up or approving a project. Similarly, an evaluation index system and a management procedure which are suitable for urban traffic characteristics and planning management procedure have been formed. In addition, for parking effect evaluation, Lin et al. [8] considered turnover rate, road trip speed and commercial traffic volume etc. indicators and used the analytic hierarchy process (AHP) model to evaluate the parking operation in the Shenzhen's on-street parking fees pilot area, and finally raise some suggestions for improvement. Yu et al. [9] used AHP and fuzzy evaluation to evaluate the parking problems in the commercial residential areas and proposed the criteria for evaluating commercial residential parking. Zhang et al. [10] adopted the fuzzy evaluation method, selected parking berths, average parking time and other indicators, comprehensively evaluated the management level of urban business districts, and finally made recommendations for its parking management. The investigation found that AHP and fuzzy evaluation are mainly used to evaluated traffic, in addition, less research on the evaluation of on-street parking, and hasn't formed a unified evaluation system or standards.

Therefore, in this study, the lean management concept will be applied to parking management. In addition, an on-street parking time-entropy model will be proposed, lean time management concept will be used to evaluated the effect of on-street parking and the result will be as a theoretical basis for on-street parking management.

2 Selection of Evaluation Factor Index

2.1 Factors of On-Street Parking Effectiveness

(1) Parking Fee

Parking fee is the main factor considered by drivers when choosing parking lot. At present, the common charging standards in domestic parking pricing include such as "5 yuan/h", "20 yuan/time", and so on, while the charging rates like "5.5 yuan/18 min"

are rare found in the actual parking charge. Under the background of developing intelligent parking system, the historical data of the use of on-street parking spaces will become an important basis for further management. For example, if the billing time unit of an on-street parking zone is"/30 min", and according to the historical statistics of the parking space collected, the parking time of 85–90% of the users is only 20–25 min. Obviously, setting the first charging time gradient at 20–25 min, some of the remaining 10–15% of the users would move their car as quickly as possible, so that parking spaces in the area would be managed more efficiently and the turnover rate would be improved. In addition, billing gradient, parking fee payer, billing method will also affect the driver's parking choice [11, 12]. The billing method include charging by time, charging by frequency, progressive pricing, etc.

(2) Parking Duration

Parking duration is mainly determined by the driver's parking purpose. Parking purpose is closely related to the nature of land which in the vicinity of the parking spaces. The walking distance from the parking space to the destination is another important factor considered by drivers when selecting the parking space. If the distance to long and drivers need to walk for a long time, it will not only reduce the willingness of drivers to use the parking lot, but also increase the parking space occupancy time and reduce parking turnover rate. According to *code for setting of on-street parking spaces* [13], the distance between on-street parking spaces and the destinations of service objects should no more than 200 m.

(3) Facility and Management

With the improvement and development of intelligent systems and modern facilities, such as Parking Guidance Intelligent System (PGIS), Intelligent Parking Reservation System (IPRS), variable information sign (VIS) and intelligent meter, the use and management of on-street parking spaces will become more efficient and convenient. At present, the charging method includes manual charge, meter charge, remote payment, etc. However, the degree of informatization and intelligence in different parking spaces are not same. Therefore, the parking facilities in different areas have the different impact on parking time extension.

Parking regulation is a necessary means to ensure on-street parking safety and parking order. Law enforcement way, punishment level and punishment ratio will affect the efficiency of on-street parking situation in time and space.

2.2 Evaluation System

Based on above analysis, three main influencing factors are selected and each factor have three evaluation indexes to establish on-street parking evaluation system in this research. The evaluation system is shown in Fig. 1.

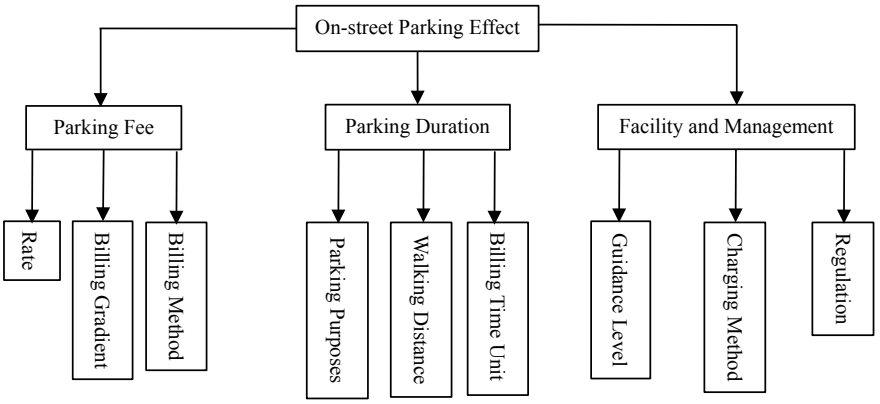


Fig. 1 Evaluation system

3 Model of On-Street Parking Time-Entropy

As a measure of uncertainty [14], entropy is widely used in the field of natural science, social science and somatology [15]. But it is seldom applied to evaluate on-street parking. The entropy weight method is an objective weighting method, which can avoid the interference of individual factors and make the evaluation results more realistic. Extended parking time used to measure factor indexes. Therefore, the entropy model was defined as on-street parking time-entropy model in this study.

Suppose there are L on-street parking spaces in a certain area Z, and the total time-entropy in Z areas is S. The time-entropy of each parking areas in Z areas is S_l, $l = 1, 2, \ldots, L$.

According to the entropy theory and the entropy weight theory [16, 17]. The equation of entropy evaluation model is:

$$S = \sum_{u=1}^{k} R_u \, S_u \tag{1}$$

u The factors of on-street parking space;
k There are k evaluation indexes under the each factor;
R_u The weight of each factor

Assuming that the evaluation model involves n on-street parking spaces and there are m evaluation indexes under one factor, the evaluation matrix of the factor is $X = (x_{ij})_{m \times n}$:

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix}_{m \times n} \tag{2}$$

x_{ij} The extended parking time of the jth parking space under the ith index

Since all indexes are measured by extended parking time in this study, it's not necessary to standardize the matrix.

The entropy of the ith index is:

$$H_i = -k \sum_{j=1}^{n} p_{ij} \cdot \ln p_{ij},\ i = 1, 2, 3, \ldots, m$$
$$k = \frac{1}{\ln n}$$
$$p_{ij} = \frac{x_{ij}}{\sum_{j=1}^{n} x_{ij}},\ i = 1, 2, 3, \cdots, m \tag{3}$$

The entropy weight of index i is:

$$r_i = \frac{1 - H_i}{m - \sum_{i=1}^{m} H_i} \tag{4}$$

Entropy weight is not used to indicate the importance of evaluation factors, but after the evaluation objects and the value of evaluation indexes are determined, it means how much effective information which the evaluation indexes provided. It is an objective and comprehensive evaluation method. The weight are determined by the amount of information passed by each index [18].

The time-entropy of factor u is:

$$S_u = \sum_{i=1}^{m} r_i H_i \tag{5}$$

The smaller the entropy of evaluation factor, the greater the variation degree of index value, the more amount of information provided, the more important role played in the comprehensive evaluation, that is, the greater weight of the index [19].

4 Example and Result Analysis

To verify the evaluation system and time-entropy model, the data of five on-street parking areas were simulated. And using the built system and model to evaluate the effect of the five parking area. The simulated data is shown in Table 1.

Taking the time-entropy calculation of the parking fee factor for example. The calculation process is as follows:

Substituting the data of the parking fee factor indexes (Table 1) into Eq. (2). The time-entropy evaluation index matrix of parking fee was obtained.

Table 1 On-street parking evaluation indexes data

Evaluation factor	Factor index	Extension parking time/min				
		Area 1	Area 2	Area 3	Area 4	Area 5
Parking fee	Rate	2.42	1.87	4.22	1.13	2.13
	Billing gradient	3.11	2.22	6.45	1.91	5.85
	Billing method	2.03	5.12	3.00	1.32	4.25
Parking duration	Parking purpose	7.32	10.71	8.20	5.71	4.3
	Walking distance	2.69	1.39	1.53	0.67	1.11
	Billing time unit	5.45	1.63	3.46	0.54	7.58
Facility and management	Guidance level	1.82	2.00	2.50	5.32	1.72
	Charging method	0.98	2.16	1.75	1.33	1.43
	Regulation	1.50	2.20	1.71	9.82	2.40

$$\boldsymbol{X} = \begin{bmatrix} 2.42 & 1.87 & 4.22 & 1.13 & 2.13 \\ 3.11 & 2.22 & 6.45 & 1.91 & 5.85 \\ 2.03 & 5.12 & 3.00 & 1.32 & 4.25 \end{bmatrix}_{3\times 5}$$

The above evaluation matrix $\boldsymbol{X}$ was substituted into Eq. (3) to obtain the time-entropy value of each index under the parking fee factor:

$$\mathrm{H}_1 = 0.9442,\ \mathrm{H}_2 = 0.9282,\ \mathrm{H}_3 = 0.9366$$

The time-entropy are substituted into Eq. (4), and the entropy weight of each index is obtained.

$$r_1 = 0.2922,\ r_2 = 0.3760,\ r_3 = 0.3318$$

Substituting the entropy weight into Eq. (5) can get the time-entropy of parking fee $S_1 = 0.9357$. The entropy weight and time-entropy of the other two factors can be obtained in the same way. The results are shown in Table 2.

Compared the time-entropy in Table 2, it shows that the time-entropy of parking fee factor is the largest, the parking duration is the second, and the facility and management time-entropy is the smallest. The result indicates that the degree of confusion caused by parking fee factor is greater than the others. Compared the entropy weight of the factor indexes, it shows that the billing time unit and regulation index have larger entropy weight. So, the more information provided by billing time unit and regulatory index. Thus, measures can be taken from these two aspects when regulators adjust on-street parking situation in this area.

Table 2 Entropy weight and time-entropy

Evaluation factor	Factor index	Entropy weight (r_i)	Time-entropy (S)
Parking fee	Rate	0.2922	0.9357
	Billing gradient	0.3760	
	Billing method	0.3318	
Parking duration	Parking purposes	0.1142	0.8799
	Walking distance	0.2467	
	Billing time unit	0.6391	
Facility and management	Guidance level	0.2379	0.8470
	Charging method	0.0740	
	Regulation	0.6881	

5 Summary

Based on the concept of lean time management, analyzed the main factors which affecting the effect of on-street parking. Established an evaluation system for on-street parking effectiveness which based on lean time management. Constructed the time-entropy model of on-street parking. Used the simulated data for evaluation. Finally, the paper puts forward the management basis for further adjustment.

Inadequacies: Only three main influencing factors have been considered to construct the evaluation system and time-entropy model in this study, and only three evaluation indexes were considered under each factor. In practice, parking is influenced by more factors. In addition, this study used simulated data rather than actual collected data. Therefore, further research and improvement is needed.

References

1. Sun J (2009) The definition of concept for total lean management. Ind Eng Manag 14(2):129–134
2. Nui ZW, Jing SW, Yang FD (2015) An analysis of driving factors of the management innovation of manufacturing enterprises based on lean management: case studies of four enterprises. Sci Sci Manage S & T 36(7):116–126
3. Li YX (2005) Research on the theory and application of lean management mode in manufacturing enterprises. Tianjin University, Tianjin
4. Dai S, Liu JG, Zhu JN et al (2014) On-street parking management strategies and practice. Urban Transp China 12(1):6–11
5. Li XM, Dai S, Liu JG (2016) The reformation and innovation of urban parking management in China. China Transp Rev 38(5):11–14
6. Alemi F, Rodier C, Drake C (2015) Cruising and on-street parking pricing: a dierence-in-dierence analysis of measured parking search time and distance in San Francisco. Transp Res Part A Policy Pract 111:187–198
7. Ma J (2016) Review on China's traffic engineering research progress: 2016. China J Highw Transp 29(6):1–161

8. Lin T, Lv GL, Tian F et al (2016) A pilot study of on-street parking charge in Shenzhen. Urban Transp China 14(4):30–39
9. Yu DY, Zhang P, Li ZQ (2009) Commercial and residential parking assessment based on analytic hierarchy process and fuzzy mathematics. China Water Transp 9(1):84–85
10. Zhang XY, Bai Y (2006) Application of fuzzy statistics approach in the parking management evaluation. Math Pract Theory 36(6):57–62
11. Guan HZ, Yan H, Li Y (2008) Travel mode choice model for parking fee payers. China Civ Eng J 41(4):91–94
12. Chen XB (2016) Pricing problem of curb parking in beijing on the basis of right of way and space-time. Beijing Jiaotong University, Beijing
13. GA/T 850-2009, Cod for setting of on-street parking spaces
14. Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27(3):379–423
15. Jia YH, Zhao J, Nan ZR et al (2006) Ecological safety assessment of grassland based on entropy-right method: a case study of Gansu pastoral area. Chin J Ecol 25(8):1003–1008
16. Niu ZW, Shi XL (2014) An effect evaluation of lean management implementation in the production system based on time entropy. Ind Eng J 17(6):17–32
17. Jia ZY, Zhao L (2010) Comprehensive evaluation of power quality based on the model of entropy weight and unascertained measure. Power Syst Prot Control 38(15):33–37
18. Feng YQ, Li XM, Li XW (2014) Comprehensive evaluation method for the railway safety based on the entropy method and the grey relation analysis. J Saf Environ 14(2):73–79
19. Ni JP, Li P, Wei CF et al (2009) Potentialities evaluation of regional land consolidation based on AHP and entropy weight method. Trans Chin Soc Agri Eng 25(5):202–209

Accurate Identification of Accident Black Point Based on Hazard Attribute Analysis

Jianyou Zhao, Yongmei Xue, Yao Peng and Chuang Zhou

Abstract After more than 30 years of rapid development, China's expressway has entered a stage of comprehensive long-term development. The safety problem of expressway has become a hotspot in the development of expressway. The accident black points as the focus of research on highway safety issues are directly related to the improvement of highway safety conditions. In this paper, the risk index is used to characterize the risk of accident multiple sections. Based on the analysis of the risk attributes of expressway accidents, the risk curve model of accident points is established. Considering the superposition of danger, Risk curve combination model, and then to promote the accident multiple point of the risk curve model, so as to establish the full range of highway safety function and curve model; to determine the accident black points to determine the threshold, combined with the highway risk curve model and determine the threshold, The establishment of highway accident black point recognition model, to achieve accurate identification of accident black points.

Keywords Expressway · Hazard attribute · Accident black point · Accurate identification

J. Zhao · Y. Xue (✉) · Y. Peng · C. Zhou
School of Automobile, Chang' an University, Xi' an 710064, China
e-mail: 1399895262@qq.com

J. Zhao
e-mail: jyzhao@chd.edu.cn

Y. Peng
e-mail: 891498699@qq.com

C. Zhou
e-mail: 851697287@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_49

1 Introduction

Now there are three categories for the study of accident black points recognition methods, one is through a simple mathematical model, these methods are simple, fast, but the accuracy is not high; one is based on statistical knowledge, these methods are objective and accurate, but the calculation is complex and these methods are lack of practicality. One is the comprehensive evaluation index method, which adopts a combination of qualitative and quantitative methods, so it has certain subjectivity.

In the identification of the accident black points, the problem has not been resolved at this time is the accident black points to identify the "Last mile" problem, this paper attempts to build a model to solve this problem.

2 Analysis on the Risk Attributes of Accident Multiple Section

In order to quantitatively analyze the degree of danger of accidents, it is necessary to analyze the attributes of accident multiple sections first.

The K20-K30 is a 10 km accident multiple section of a typical expressway, Including the accident black point (K24-K26) and non-accident black point. The statistical distribution of the accident is shown in Table 1, and the statistics use the compound statistical approach.

K20-K30 segment accident statistics distribution is shown in Fig. 1. In order to study the accident distribution of K20-K30, the K20-K30 accident distribution is verified by normal function. The verification results are shown in Table 2.

The number of accidents normal P-P diagram consistent with the regularity of normal distribution (Fig. 2).

In the accident multiple section K20-K30, the law of accident occurrence is similar to the law of normal distribution.

In order to quantitatively analyze the extent of the risk of the accident, make the following definition: W represents the danger of the points on the road, and the value of W is positively related to the risk of the points on the road. The cumulative mileage

Table 1 K20-K30 accident statistics

Starting point	Terminal point	Number of accidents	Starting point	Terminal point	Number of accidents
20	22	0	25	27	4
21	23	1	26	28	2
22	24	2	27	29	1
23	25	5	28	30	0
24	26	6			

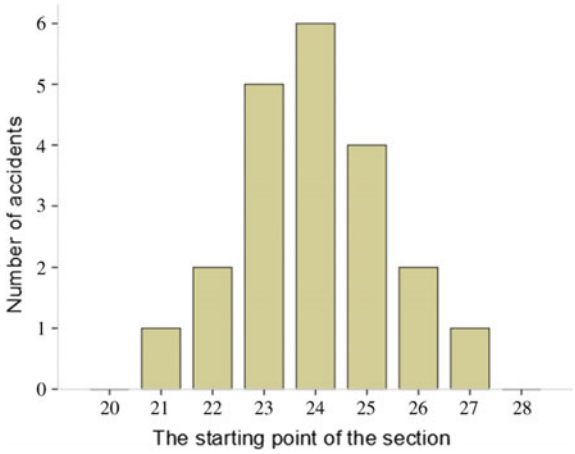


Fig. 1 K20-K30 section of the accident statistics distribution diagram

Table 2 Normal model description

The name of the model		MOD_2
Non-seasonal difference		0
Seasonal difference		0
The length of seasonal period		No periodicity
Distribution	Classification	Normal
	Position	Estimate
	Scale	Estimate

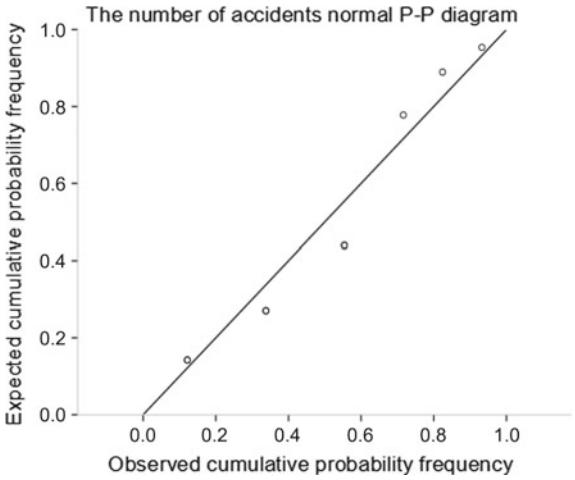
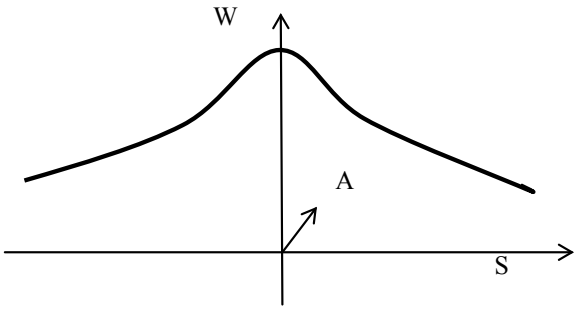


Fig. 2 The number of accidents normal P-P diagram

Fig. 3 Road risk curve



S is the abscissa, the risk of (w) is the ordinate, and zero is the accident-prone point. The risk of a road as is shown in Fig. 3.

Select the normal function to characterize the road risk, its rationality lies in:

(1) The random variable W in the normal curve reaches the maximum at the point where the accident occurs, and then decreases toward both sides. To characterize the road risk with normal function, consistent with the actual situation of road continuity and risk experience.
(2) In natural and social phenomena, a large number of random variables are obeyed or approximately obey the normal distribution.
(3) The central limit theorem. Since many random variables are formed by the combined effects of a large number of independent stochastic factors. if each of the single factors in the overall effect play a small role, then the random variable is often similar to the normal distribution.

Therefore, the normal distribution curve as a black dot model has a sufficient theoretical and practical basis [1].

3 Analysis of Hazard Curve Model of Accident Point

It is determined that the risk distribution of road sections conforms to the normal function, and then the risk curve of the accident occurrence point is modeled and the parameters are determined. The random variable W follows the normal distribution, and its probability density function is as follows: (1):

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\sigma > 0, -\infty < x < +\infty) \tag{1}$$

The standard normal distribution is given by (2):

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (-\infty < x < +\infty) \tag{2}$$

Generally there are two parameters in the normal function, that is (μ, σ), the following μ and σ are discussed separately.

3.1 Parameter μ

In the normal function, the mathematical properties of the parameter μ represent the mean of the function, and the parametric image characteristic of the parameter μ represents the axis of symmetry of the normal function. In this paper, the parameter μ can determine the specific location where road risk curve is located on the road by determining the location of the accident point. Define L is the location of the accident where the risk of the road is the highest.

3.2 Parameter

In the normal curve, the larger the parameter σ, the greater the discretization of the random variable, the more gentle the curve is. The smaller the parameter σ, the smaller the discretization of the random variable, the steeper the curve is, and the value of the random variable is closer to the mean μ, the greater the f(x). Curve graphics' change with the parameter σ as is shown in Fig. 4.

Since the value of x represents the discrete case of the distribution of the accident, the following assumptions are made:

① Assuming that σ is negatively correlated with the severity of the accident, then when the value of σ is small, the maximum value of the curve should be larger and the overall picture should be relatively concentrated, and the curve should be located above the curve with the larger value of σ relative to the general event; ② Assuming that σ is positively related to the severity of the accident, the larger the value of σ, the greater the peak of the curve should be.

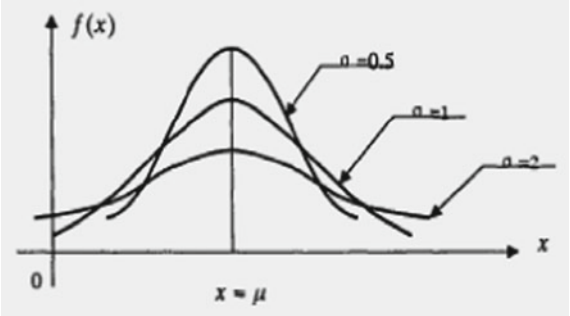


Fig. 4 The change of the normal function with σ

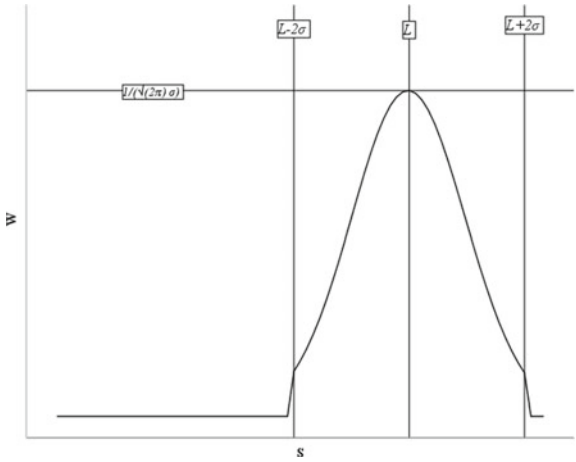


Fig. 5 Schematic diagram of the risk curve

The above assumptions are clearly not consistent with Fig. 4, so the value of x can not be used as a measure parameter.

Since σ can not characterize the nature of the accident, then the "3 σ principle" can be used for research. "3 σ principle" is the probability that the values are distributed in $(\mu - \sigma, \mu + \sigma)$, $(\mu - 2\sigma, \mu + 2\sigma)$, $(\mu - 3\sigma, \mu + 3\sigma)$ are 0.6826, 0.9544, 0.9974.

The value of W is almost always concentrated in the range of $(\mu - 3\sigma, \mu + 3\sigma)$, and the probability of exceeding this range is less than 0.3%. However, it is noted that the probability of curve on the interval $(\mu - 2\sigma, \mu + 2\sigma)$ is 0.954, and curve length reduced by nearly 1/3 relative to interval $(\mu - 3\sigma, \mu + 3\sigma)$, which can meet the requirements of traffic engineering, the part whose probability is too small can be neglected reasonably [2].

According to the above analysis, the risk curve of the accident can be modeled as follows: (3):

$$W_i = f(x_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \tag{3}$$

where $\mu = L$, $L - 2\sigma \leq x_i \leq L + 2\sigma$

The risk diagram is shown in Fig. 5.

3.3 Hazard Difference Variable

There is no general uniform definition of the current black spots. In order to reflect the difference in the degree of risk, We need select an indicator, whose changes can distinguish the severity of the accident. For this reason A is introduced by multiplying A by the standard normal distribution function as the final distribution function.

Different values of A, the curve shape is shown in Fig. 6.

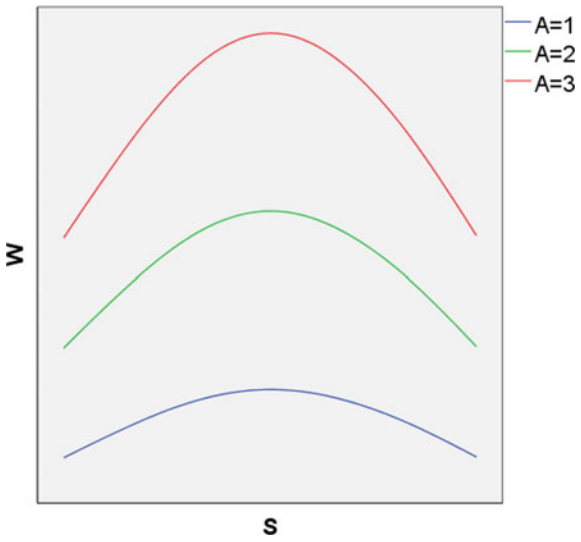


Fig. 6 Risk curves for different A values

It can be seen from Fig. 6, different values of A, the different shapes which the curve shows can distinguish the degree of danger of the road commendably. And each point on the risk index for a high-risk section is grater than a general dangerous road.

The modified risk function is given by (4):

$$W_i = \frac{A_i}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - L)^2}{2\sigma^2}} \tag{4}$$

Among them $L - 2\sigma \leq x_i \leq L + 2\sigma$

As the value of A is to more rationally characterize the severity of the accident, the selection for value of A can learn from the equivalent number of accidents, such as (5):

$$A_i = Q_i + \alpha D_i + \beta J_i + \gamma M_i \tag{5}$$

where:

Q Is the total number of accidents at that point;
D Is the total number of deaths in the accident;
J Is the number of injuries in the accident;
M Is the property loss in the accident;
α Is the weight of death accident;
β Is the weight of injured accident;
γ Is the Weight of loss for property.

According to the research results of domestic and foreign to determine the weight of death accident 2.0, the weight of injured accident 1.5, the weight of loss for property 1.

That is: $\alpha = 2.0$, $\beta = 1.2$, $\gamma = 1$.

Than is [1]:

$$A_i = Q_i + 2.0D_i + 1.5J_i + M_i \tag{6}$$

3.4 Accident Point Risk Curve Function

From 2.1, 2.2, 2.3 can get the accident point risk curve function as shown in Eq. (7):

$$W_i = \frac{Q_i + 2.0D_i + 1.5J_i + M_i}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x_i - L)^2}{2\sigma^2}} \tag{7}$$

Among them $L - 2\sigma \leq x_i \leq L + 2\sigma$

4 Analysis on the Hazard Curve Model of Accident Black Spot

4.1 Two Accident Point Risk Curve Model

The risk curve model of a single accident is known. The risk curve at the accident black spot is superimposed by the risk curve of different single accident. First analyze the risk curve composed of two accident points. The simple model is shown in Fig. 7.

There are overlapping parts of the two risk curves, we can see that the risk is at a high level between the two accident points. Its risk is affected by two accident points, the index of its risk can be set as a superposition of both, namely:

$$W_j = \sum \left(\frac{Q_i + 2.0D_i + 1.5J_i + M_i}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x_j - L)^2}{2\sigma^2}} \right) \tag{8}$$

Among them i = 1, 2, $L_1 - 2\sigma \leq x_i \leq L_2 + 2\sigma$

The adjusted risk profile of the two accident points is shown in Fig. 8.

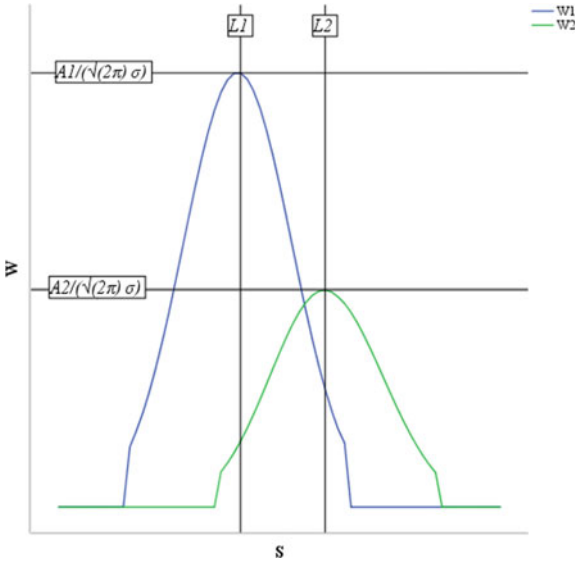


Fig. 7 A simple model of the risk curve of two accident points

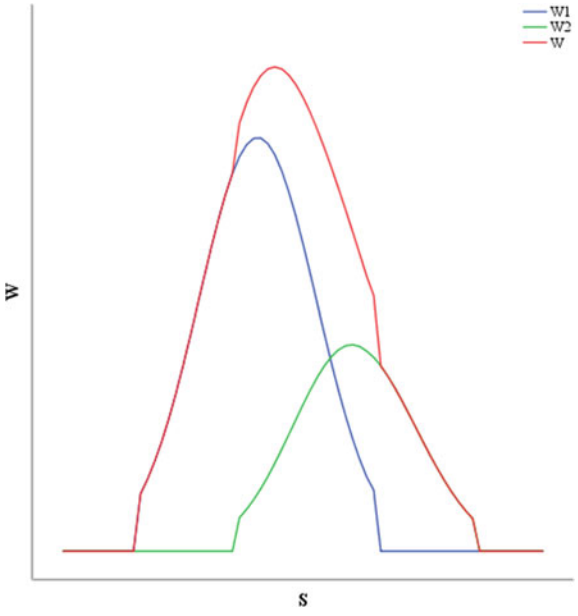


Fig. 8 Risk curve adjustment model of two accident points

4.2 The Value of σ

The distribution curve of the normal distribution is shown on the 10 and 1 km sections of Luoluan Expressway. In this paper, the statistical standard used is 2 km cumulative

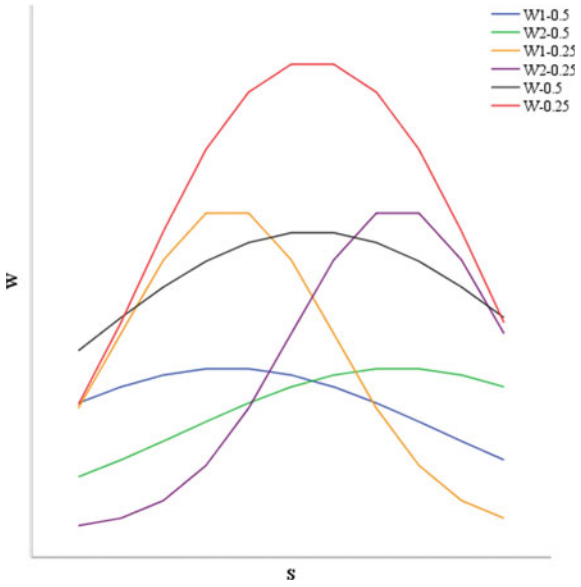


Fig. 9 Compare graph of the value of σ

stack in the statistical accident data of Luoluan Expressway, and the range of the risk curve is included in the standard step of 2 km. 2, 1 km can be selected as the range of the risk curve. That is:

$$4\sigma = 1, \sigma = 1/4 \text{ or } 4\sigma = 2, \sigma = 1/2$$

Respectively, $\sigma = 1/4$ and $\sigma = 1/2$, as is shown in Fig. 9.

It can be seen that the trend of the risk curve of the accident road can be more obvious when $\sigma = 1/4$ than $\sigma = 1/2$, and the accident point can be clearly identified. When $\sigma = 1/2$ the whole curve is relatively gently, and the trend is not prominent.

In summary can be learned: $\sigma = 1/4$ is more appropriate value.

4.3 Multiple Accident Point Risk Curve Model

The hazard curve function obtained from $\sigma = 1/4$ is shown in Eq. (9):

$$W_i = \frac{4(Q_i + 2.0D_i + 1.5J_i + M_i)}{\sqrt{2\pi}} e^{-8(x_i - L)^2} \tag{9}$$

Among them $L - 0.5 \leq x_i \leq L + 0.5$

The risk curves of two accident points in (8) are extended. When the number of accident points in the section is n, the risk curve function of accident multiple point is shown by Eq. (10):

$$W_j = \sum \left(\frac{4(Q_i + 2.0D_i + 1.5J_i + M_i)}{\sqrt{2\pi}} \mathrm{e}^{-8(x_j - L)^2} \right) \tag{10}$$

Among them i = 1, 2, …, n, $L_1 - 0.5 \leq x_j \leq L_n + 0.5$

4.4 The Risk Function of Full Section of the Highway

For a given statistical period, any waypoint on the expressway can be divided into two categories: there was no accident in the range of 0.5 km before and after the point and the occurrence of n accidents in the range of 0.5 km before and after the point.

It is known that the risk function of whole expressway is given by (11):

$$W_j = \begin{cases} \sum \left(\frac{4(Q_i + 2.0D_i + 1.5J_i + M_i)}{\sqrt{2\pi}} \mathrm{e}^{-8(x_j - L)^2} \right), n > 0 \\ 0, n = 0 \end{cases} \tag{11}$$

Among them i = 1, 2, …, n, $L_1 - 0.5 \leq x_j \leq L_n + 0.5$

4.5 The Discrimination Threshold of Accident Black Point

To measure a curve, we often use the characteristic quantity: average value $\overline{W}$ and standard deviation σ.

(1) The average value $\overline{W}$ is the risk level of the risk curve as a whole, which reflects the magnitude of the risk of the section, and the average calculated by the formula (12):

$$\overline{W} = \frac{w_1 + w_2 + \ldots + w_n}{n} = \frac{\sum_{i=1}^{n} w_i}{n} \tag{12}$$

(2) is an risk range indicator that the risk profile as a whole covers, which reflects the degree of risk polymerization of the section. The standard deviation is calculated from Eq. (13):

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2} \tag{13}$$

The discrimination threshold of accident black point is:

$$\overline{W} > \overline{W}_{blackpoint}$$

$$\sigma > \sigma_{blackpoint}$$

5 Example Application

Luo Luan Expressway is divided into Luoyang to Songxian section and Songxian to Luanchuan section, the length of route is 125 km, the width of road is 28 m, the design speed is 100 km/h. Including 1111 m grand bridge, the total length of medium bridge is 45,807 m, the length of great bridge is 952 m. With 131 channels Culverts 131 channels, 52 channel, 26 tunnels, the total length of the tunnel is more than 20 km.

Known Luoluan Expressway accident black spots and the road conditions at accident black points as is shown in Table 3, meanwhile analysing the risk curve of different special sections and general sections.

We get the accident black points curve of the special road section and the general road, and get the average value and standard deviation as is shown in Table 4.

Table 3 Luo Luan expressway accident black points and road conditions table

Road classification	Accident black point	Main road conditions	
Bridge section	K3-K5	Qianxi River Bridge	Liang Liu Bridge
	K17-K19	Louzigou Bridge	Tangwa Bridge
General section	K24-K26	General section	
Curved section	K33-K35	Curved combination	
Bridge bend section	K38-K40	Luhun Reservoir Main Canal Bridge	Curved combination
	K46-K48	Tiger ditch Bridge	Curved combination
Tunnel section	K80-K82	Wudao Temple No. 1.2 Tunnel	Liu Ping No. 1.2.3 tunnel

Table 4 Luo Luan expressway accident black point threshold value of distinguishing

Road classification	Accident black point	Average value of road sections	Standard deviation of road sections σ	Average value of road conditions	Standard deviation of road conditions
Bridge section	K3-K5	6.3598	3.1478	7.0189	3.4082
	K17-K19	7.8714	3.7451		
Tunnel section	K80-K82	11.8477	7.9139	11.8477	7.9139
Curved section	K33-K35	4.4362	3.3351	4.4362	3.3351
Bridge bend section	K38-K40	22.1961	15.4741	18.0150	12.8222
	K46-K48	12.2038	9.1363		
General section	K24-K26	4.0435	2.9092	4.0435	2.9092

The value of the accident point and the accident point equivalent (i.e., the value of the accident point A) of the whole section of Luo Luan Expressway are shown in Table 5.

The risk curve of the whole section of Luo Luan Expressway is shown in Fig. 10.

Accurate identification of accident black spots, the risk average 4.043 is the minimum risk average of all sections, and it is used as a benchmark to get screening sections. That is, the risk coefficient is greater than or equal to the road section of

Table 5 Luo Luan expressway accident point equivalent value

Accident point	Equivalent number of accidents	Accident point	Equivalent number of accidents	Accident point	Equivalent number of accidents
3.038	2.44	29.9	1.35	58.4	1.571
3.23	4.49	29.95	5.74	61.15	4.55
4.4	6.28	31	3.50955	61.815	2.07
5.6	5.83	31.85	1.575	64.1	2.05
6.65	4.2725	33.5	4.644	64.3	2.7
7.1	1.944	33.6	1.715	65.4	1.28
7.3	1.2705	33.9	1.125	70.25	1.315
8.3	1.7925	34.4	1.362	80.1	11.245
10.3	1.25	35.53	2.12	80.3	1.51
10.7	2.21	36.8	3.39	80.5	2.85
11.6	1.37	37.42	2.285	80.6	1.6105
13.3	3.01	38.84	2.1775	83.5	2.02
14.5	2.81	38.95	6.6	93.96	1.4
18.3	3.39	39.05	20.481	95.5	3.0025
18.45	2.54	39.1	1.8175	98.1	1.5929
18.65	1.55	39.5	3.73	98.3	3.8175
18.8	2.73	39.8	4.85	100.1	1.3305
19.3	1.69	46.1	1.33	101.8	1.56
21.3	1.93	46.3	17.36	103.4	1.206
23.2	3.5685	46.55	1.22	103.8	1.37
23.6	1.73	47.2	1.25	105.4	1.6475
24.06	4.195	47.5	5.365	106.4	3.88
24.3	2.2	47.7	1.575	107.25	1.705
25.1	1.16	50.05	1.52	109.9	6.5075
25.9	1.655	50.5	2.2425	111.7	1.5175
25.9	1.345	51.3	1.615	112.7	1.34
26.2	1.225	56.5	1.305	114.85	2.355
27.4	1.12	58.2	1.325	122.85	1.8784

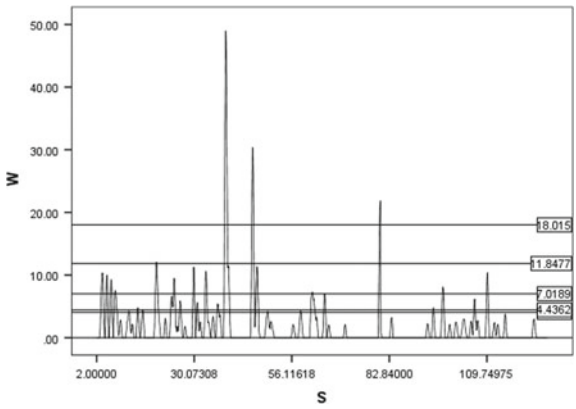


Fig. 10 Luoluan expressway accident black points accurate identification diagram

general risk coefficient. It can be seen that the risk factors under special road conditions are included in the road to be screened. The screening results are shown in Table 6.

The accurate identification illustration of the accident black points is shown in Fig. 10.

You can get accurate identification of the incident black points:

Bridge section: K2 + 950-K3 + 380, K4 + 210-K4 + 600, K5 + 440-K5 + 790, K6 + 620-K6 + 860, K98 + 110-K98 + 380, K109 + 690-K110 + 100;
Tunnel section: K79 + 870-K80 + 300;
Curved section: K13 + 210-K14 + 520, K18 + 060-K19 + 100, K33 + 220-K33 + 950;
Bridge bend section: K38 +660-K39 + 460, K46 + 050-K46 + 560;
General section: K23 + 020-K24 + 470, K25 + 740-K26 + 220, K29 + 610-K31 + 180, K58 + 240-K58 + 390, K63 + 950-K64 + 500.

6 Conclusion

Based on the analysis of the risk attributes of the accident black spots, the paper establishes the hazard descriptive model of the accident point, and then promotes the accident multiple point hazard description model, and finally establishes the curve model of the accident black point based on the risk attribute and analyze and determine the parameter. The above model is applied to the Luo Luan Expressway for example verification, and get black points which were identified accurately of the Luo Luan Expressway.

Table 6 Accurate identification results of Luo Luan expressway accident black point

The road to be screened	Road condition	4.0435	4.4362	7.0189	11.8477	18.0150
K2 + 810-K3 + 510	1			K2 + 950-K3 + 380		
K4 + 070-K4 + 710	1			K4 + 210-K4 + 600		
K5 + 300-K5 + 920	1			K5 + 440-K5 + 790		
K6 + 420-K7 + 330	1			K6 + 620-K6 + 860		
K36 + 640-K37 + 140	1			0		
K95 + 380-K95 + 620	1			0		
K97 + 940-98 + 560	1			K98 + 110-K98 + 380		
K109 + 58-K110 + 230	1			K109 + 690-K110 + 100		
K38 + 530-K40 + 080	4					K38 + 660-K39 + 460
K45 + 810-K46 + 770	4					K46 + 050-K46 + 560
K47 + 090-K47 + 880	4					0
K60 + 890-K61 + 520	2				0	
K79 + 670-K80 + 300	2				K79 + 870-K80 + 300	
K106 + 190-K106 + 610	2				0	
K10 + 530-K10 + 690	3					
K13 + 160-K14 + 600	3		K13 + 210-K14 + 520			
K18 + 060-K19 + 130	3		K18 + 060-K19 + 100			
K50 + 300-K50 + 490	3		0			
K33 + 190-K33 + 980	3		K33 + 220-K33 + 950			
K23 + 020-K24 + 470	5	K23 + 020-K24 + 470				

(continued)

Table 6 (continued)

The road to be screened	Road condition	4.0435	4.4362	7.0189	11.8477	18.0150
K25 + 740-K26 + 220	5	K25 + 740-K26 + 220				
K29 + 610-K31 + 180	5	K29 + 610-K31 + 180				
K58 + 240-K58 + 390	5	K58 + 240-K58 + 390				
K63 + 950-K64 + 500	5	K63 + 950-K64 + 500				

Note 1 The road conditions to be screened are: 1, 2, 3, 4, 5 on behalf of the bridge section, tunnel section, curved section, bridge bend section, the general section

Note 2 The risk index for the screening section is greater than the baseline 4.0435. The screening process takes K2 + 81-K3 + 510 as an example, the K2 + 81-K3 + 510 segment is the bridge section. See Table 4. The average of judgment threshold of the bridge segment is 7.0189. Screening the section whose hazard index is greater than 7.0189 from the K2 + 81-K3 + 510 section, that is, K2 + 950-K3 + 380 section, and so on

Note 3 0 indicates that there is no accident black point section that meets the road condition baseline in the section

References

1. Liu Y (2006) Study on the intelligence batch-check and countermeasures of traffic accident black-spots. Ph.D.'s degree thesis of Southwest Jiaotong University, Chengdu
2. Huang B (2009) Research on identification hazard locations and analysis ITS causes of accident in freeway of mountain area. Master's degree thesis of Harbin University of Technology, Harbin

Spatio-Temporal Autocorrelation-Based Clustering Analysis for Traffic Condition: A Case Study of Road Network in Beijing

Wei Wei, Qiyuan Peng, Ling Liu, Jun Liu, Bo Zhang and Cheng Han

Abstract Traffic congestion is an increasingly serious problem worldwide. In the last decade, many cities have paid great efforts to establish Intelligent Transportation Systems (ITS), and a large amount of spatio-temporal data from traffic monitoring system is also accumulated. However, with the devices and facilities of ITS getting completed, effectiveness of ITS practices is always restricted by traffic information fusion and exaction technique. Traffic condition-determining is a crucial issue for Advanced Traffic Management Systems, on which many researchers have done profound studies. The existing studies are mostly focused on traffic condition recognition at a certain road and time point; while in practice, it's more meaningful how different kinds of traffic condition are correlated and distributed in space-time. Therefore, in this research we present an improved spatio-temporal Moran scatterplot (STMS), by which traffic conditions are pre-classified into four types: homogenous uncongested traffic, heterogeneous uncongested traffic, homogenous congested traffic and heterogeneous congested traffic. Then at the basis of STMS, a novel spatio-temporal clustering method combining pre-classification of traffic condition is proposed. Finally, the feasibility and effectiveness of the clustering methodology are demonstrated by case studies of Beijing. Result shows that the proposed clustering method can not

W. Wei (✉) · Q. Peng
School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China
e-mail: 815612439@qq.com

Q. Peng
e-mail: qiyuan-peng@swjtu.edu.cn

W. Wei · L. Liu · J. Liu · B. Zhang · C. Han
Beijing National Railway Research & Design Institute of Signal & Communication Group Co., Ltd., Beijing 100073, China
e-mail: 627ll@crscd.com.cn

J. Liu
e-mail: junliu@crscd.com.cn

B. Zhang
e-mail: zhangbo@crscd.com.cn

C. Han
e-mail: hc@crscd.com.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_50

only effectively reveal the relation of traffic demand to road network facilities, but also recognize the road sections where congestion originates or gets alleviated in the network, which provides foundations for traffic managers to alleviate congestion and improve urban transport services.

Keywords Spatio-temporal clustering · Traffic condition · STMS · Spatio-temporal autocorrelation

1 Introduction

Traffic congestion is an increasing serious problem in many metropolises all over the world. For the instance of Beijing, as the number of automobiles approaching 5.4 million in 2014, the average speeds on urban expressways are just 37.1 km/h and 32.2 km/h for the early and the late peak respectively [1], which are significantly lower than the design speed (80 km/h).

In order to effectively alleviate traffic congestion, many cities have paid large efforts to establish Intelligent Transportation Systems (ITS) in the last decade. After years of operation, a large amount of traffic data is accumulated from loop detectors, Global Positioning Systems (GPS) devices or Remote Traffic Microwave Sensors (RTMS). Such traffic data can be seen as spatio-temporal series, which is a set of continuous observations of traffic flow on roads correlated with each in the network [2–5].

However, with the devices and facilities of ITS getting mature and completed, the lack of effective methods on information fusion and exaction based on traffic data restricts the development and effectiveness of ITS. The success of ITS is very dependent on the quality of traffic information, and one of the critical needs is to recognize traffic condition [6]. Especially, traffic condition identification is also a crucial foundation to improve the managerial level and effectiveness for traffic management departments.

Most traffic operation centers classify traffic conditions into congested traffic and uncongested traffic, while congestion can be further divided into two types by many researchers: recurrent congestion and non-recurrent congestion [7–9]. According to the physical characteristics, traffic conditions can also be classified into three groups: free flow, wide moving jam and synchronized flow [10, 11]. As traffic congestion usually spreads from a certain road to others in the network, it's more meaningful for traffic managers how traffic conditions on different roads are correlated and distributed in space and time than what the traffic condition it is at a certain location and time.

Spatio-temporal object whose characteristics are various over space and time has got extensive applications in space and time study [12]. In spatial econometrics, it is thought that the properties of a spatio-temporal object are always related to that of the others adjacent to it in space-time, which can be called spatio-temporal autocorrelation and measured by indicators such as space-time autocorrelation index

[13], spatio-temporal Moran's I [14, 15], Spatio-temporal autocorrelation and partial autocorrelation functions [16]; Kamarianakis and Prastacos [17]. So, the spatio-temporal autocorrelation analysis of traffic can provide an insight into the correlative structure of urban traffic condition.

Spatio-temporal clustering aims at recognizing the groups of spatio-temporal objects with similar characteristics and features [18]. The spatio-temporal clustering algorithm mainly falls into three types: scan statistics method [19, 20], density-based method [21, 22] and distance-based method (Kulldrff and Hialmars [23, 24]. There are widespread applications of spatio-temporal clustering in many fields such as public health and safety [19, 20], earthquake monitoring [21, 22] and climate change [25, 26].

The purpose of this paper is to develop a spatio-temporal clustering method to study the distribution of different traffic condition. As traffic conditions on different roads are generally correlated in space and time, we use an improved spatio-temporal Moran scatterplot (STMS) to make a classification of urban traffic conditions firstly. A novel spatio-temporal clustering method combining pre-classified with STMS is then developed. Finally, the case studies of Beijing are conducted to demonstrate the feasibility and effectiveness of proposed method.

2 Urban Traffic Condition Classification

Generally, traffic congestions are not local or static phenomena, they do spread from a certain road to others in the network. Therefore, for traffic managers, it's more meaningful how traffic conditions on different roads are correlated in space and time than what the traffic condition it is at a certain location and time. In this section, an improved spatio-temporal Moran scatterplot (STMS) is proposed to explore the local spatio-temporal correlation of urban traffic, and further to make a classification.

2.1 Improved Spatio-Temporal Moran Scatterplot

Moran's I is proposed by Moran in [27] to study spatial autocorrelation, which has got extensive application with the contribution of researchers latter [14, 15, 28–31]. As a local analysis method of spatial autocorrelation, Moran scatterplot (MS) was derived from Moran's I by Anselin in [32]. We extend spatial MS to the aspect of space-time, called spatio-temporal Moran scatterplot (STMS), to study the local spatio-temporal correlation of urban traffic conditions.

The condition at a specific spatial unit p and a particular time point i can be defined as a spatio-temporal object $ST_{(p,i)}$. According to traditional MS, $Wz_{(p,i)}$ and $Z_{(p,i)}$ in the improved STMS can be formulated as follows.

$$Wz_{(p,i)} = \frac{\sum_{q=0}^{N}\sum_{j=0}^{T} w_{(p,i)(q,j)} Z_{(q,j)}}{\sum_{q=0}^{N}\sum_{j=0}^{T} w_{(p,i)(q,j)}} \tag{1}$$

$$Z_{(p,i)} = \frac{(y_{(p,i)} - \bar{y})}{\sqrt{\frac{\sum_{p=0}^{N}\sum_{i=0}^{T}(y_{(p,i)} - \bar{y})^2}{NT-1}}} \tag{2}$$

where $Z_{(p,i)}$ is the standardized attribute value of $ST_{(p,i)}$, and $Wz_{(p,i)}$ is the weighted standardized attribute value of its spatio-temporal neighbors. N and T define the number of spatial units and time points studied respectively. $w_{(p,i)(q,j)}$ is a binary variable which denotes the spatio-temporal adjacent relationship between spatio-temporal object $ST_{(p,i)}$ and $ST_{(q,j)}$. $y_{(p,i)}$ represents the studied attribute of $ST_{(p,i)}$ and $\bar{y}$ indicates the average value of all spatio-temporal objects.

Using $Z_{(p,i)}$ and $Wz_{(p,i)}$ as abscissa and ordinate respectively, spatio-temporal object can be mapped to a point in this two-dimensional coordinates plane, which forms the spatio-temporal Moran scatterplot (STMS). In STMS, point located in the first quadrant means that the studied attribute value of both spatio-temporal object and its neighbors are significantly higher than the average, and point in the third quadrant means that spatio-temporal object and its neighbors are both relatively lower than the average level in studied attribute. The second quadrant shows that spatio-temporal object takes low value of studied attribute with its neighbors having high values, while the forth quadrant illustrates the opposite.

2.2 Traffic Condition Classification with STMS

The speed of traffic flow can quantitatively reflect the performance of urban road traffic, and is strongly related to travel time at a certain road section. Compared with the travel time data that are relatively difficult to be directly measured in a large-scale network, speed data are readily collected by monitoring system [6]. Considering the variance of travel speed on different type of road, a relative speed through the hundredfold ratio of the actual speed to design speed is used as indicator to evaluate traffic condition.

For urban traffic, spatio-temporal object can be defined as the traffic condition at a certain road and time. Taking the relative speed ratio as the studied attribute of spatio-temporal object, traffic condition can be preliminarily divided into two types: congested traffic and uncongested traffic. According to STMS, traffic conditions can be further divided into 4 types as shown in Table 1.

To further explain the classification above, the distribution of 4 kinds of traffic condition for the Second Ring Road in Beijing are shown as Fig. 1. As seen in Fig. 1, the homogenous uncongested traffic clusters at the midnight. While during the peak hours, there are mainly spatio-temporal objects with homogenous congested traffic.

Table 1 Traffic condition classification with STMS

Traffic condition classification	Location in STMS	Characteristics
Homogenous uncongested traffic	First quadrant	Spatio-temporal object and its neighbors both with uncongested traffic
Heterogeneous uncongested traffic	Second quadrant	Spatio-temporal object with congested traffic; its neighbours with uncongested traffic
Homogenous congested traffic	Third quadrant	Spatio-temporal object and its neighbors both with congested traffic
Heterogeneous congested traffic	Fourth quadrant	Spatio-temporal object with uncongested traffic; its neighbours with congested traffic

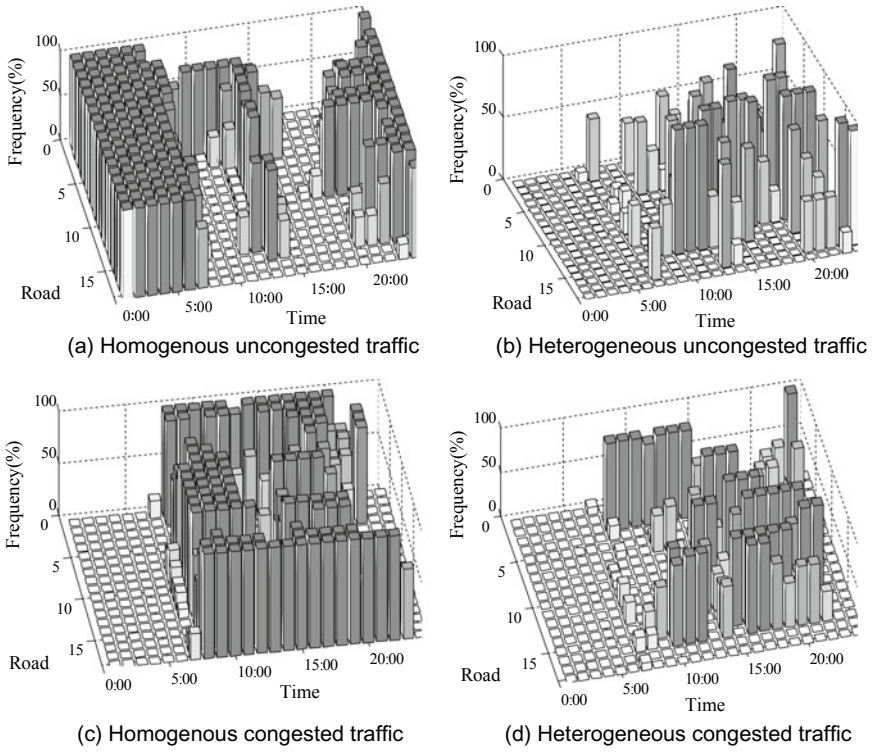


Fig. 1 Distribution of 4 kinds of traffic condition of the second ring road in Beijing

In the off-peak hours of the day, heterogeneous uncongested and congested traffic are majority. Roads with heterogeneous uncongested traffic tend to become congested earlier and get uncongested later than its neighbours during peak hours, which are bottlenecks in the network. While roads with heterogeneous congested traffic become congested later and get uncongested earlier, which play the role of dredging congestion in the network. So, the STMS can effectively reveal the spatio-temporal correlation structure of traffic condition and recognize road where congestion originates or gets alleviated in the road network.

3 Spatio-Temporal Clustering Method Combining Pre-classification

Spatio-temporal objects with the same kind of traffic condition tend to aggregate in space and time in Fig. 1, so we can use spatio-temporal clustering to further analyze the distribution of urban traffic. According to STMS, as each kind of traffic condition has distinct characteristics, it is more reasonable to take traffic condition classification as clustering basis rather than the raw traffic data. In this section, an automatic clustering method combining pre-classification with STMS is proposed.

3.1 Pre-classification of Traffic Condition

On the basis of STMS, traffic conditions can be pre-classified into four classes: homogenous uncongested traffic, heterogeneous uncongested traffic, homogenous congested traffic and heterogeneous congested traffic, represented by the integer value from 1 to 4 of $m_{(p,i)}$ respectively:

$$m_{(p,i)} = \begin{cases} 1 \; Z_{(p,i)} \geq 0 \, and \; Wz_{(p,i)} \geq 0 \\ 2 \; Z_{(p,i)} < 0 \, and \; Wz_{(p,i)} \geq 0 \\ 3 \; Z_{(p,i)} < 0 \, and \; Wz_{(p,i)} < 0 \\ 4 \; Z_{(p,i)} \geq 0 \, and \; Wz_{(p,i)} < 0 \end{cases} \tag{3}$$

For the purpose of making adjacent spatio-temporal objects with similar characteristics one cluster, the distance in similarity sense between pair of spatio-temporal objects should be computed firstly. For $ST_{(p,i)}$ and $ST_{(q,j)}$, their distance $d_{(p,i)(q,j)}$ can be calculated as Eq. (4).

$$\begin{aligned} d_{(p,i)(q,j)} &= 1 - w_{(p,i)(q,j)} \times \delta_k(m_{(p,i)}, m_{(q,j)}) \\ \delta_k(m_{(p,i)}, m_{(q,j)}) &= \begin{cases} 1 \; if \; m_{(p,i)} = m_{(q,j)} \\ 0 \; else \end{cases} \end{aligned} \tag{4}$$

In Eq. (4), $d_{(p,i)(q,j)}$ is a binary variable because $w_{(p,i)(q,j)}$ and $\delta_k(m_{(p,i)}, m_{(q,j)})$ are both binary. $w_{(p,i)(q,j)}$ represents the spatio-temporal adjacent relation between $ST_{(p,i)}$ and $ST_{(q,j)}$ as mentioned earlier, while $\delta_k(m_{(p,i)}, m_{(q,j)})$ denotes the consistency of traffic condition classification between $ST_{(p,i)}$ and $ST_{(q,j)}$.

3.2 Spatio-Temperal Clustering Method

To facilitate clustering calculation, the logical operations: spatio-temporal link and spatio-temporal subordination are introduced.

Spatio-temporal link: if $d_{(p,i)(q,j)} = 0$, spatio-temporal object $ST_{(p,i)}$ and $ST_{(q,j)}$ are spatio-temporally linked to each other, represented by $ST_{(p,i)} \leftrightarrow ST_{(q,j)}$.

Spatio-temporal subordination: For spatio-temporal object $ST_{(p,i)}$ and a collection SP_Set of spatio-temporal objects ($ST_{(p,i)} \notin SP\_Set$), if there is any spatio-temporal object $ST_{(q,j)}$ in SP_Set satisfying $ST_{(p,i)} \leftrightarrow ST_{(q,j)}$, then $ST_{(p,i)}$ subordinates to SP_Set, denoted by $ST_{(p,i)} \rightarrow SP\_Set$.

Meanwhile, in the process of clustering, the spatio-temporal object chosen as the first element of a new cluster is called spatio-temporal kernel. Not all spatio-temporal objects can be chosen as spatio-temporal kernel, and only those spatio-temporally linked to a larger number of spatio-temporal objects than a certain lower limit value ε can possibly become spatio-temporal kernel, called candidate spatio-temporal kernel.

3.2.1 Algorithm Flowchart of Spatio-Temporal Clustering

The automatic clustering process combining pre-classification with STMS is schematized as Fig. 2, and the main steps are as follows.

Step 1: Determine the pre-classification value $m_{(p,i)}$ of every spatio-temporal object based on STMS.
Step 2: Calculate the distance $d_{(p,i)(q,j)}$ in similarity sense between each pair of spatio-temporal objects.
Step 3: Build the distance matrix of all spatio-temporal objects with $d_{(p,i)(q,j)}$, and count the number of other spatio-temporal objects each spatio-temporal object $ST_{(p,i)}$ spatio-temporally linked to, denoted by $Ln_{(p,i)}$.
Step 4: Find the largest $Ln_{(p,i)}$ value $\max(Ln_{(p,i)})$, and compare it to the threshold value ε. There will be two conditions:

(1) If $\max(Ln_{(p,i)})$ is greater than or equals ε, set the spatio-temporal object (or any one of spatio-temporal objects) with the largest $Ln_{(p,i)}$ value as the spatio-temporal kernel of a new spatio-temporal cluster SP_Set. And then, add all spatio-temporal objects spatio-temporally subordinating to the new cluster in it recurrently. Restart this step to construct another cluster.

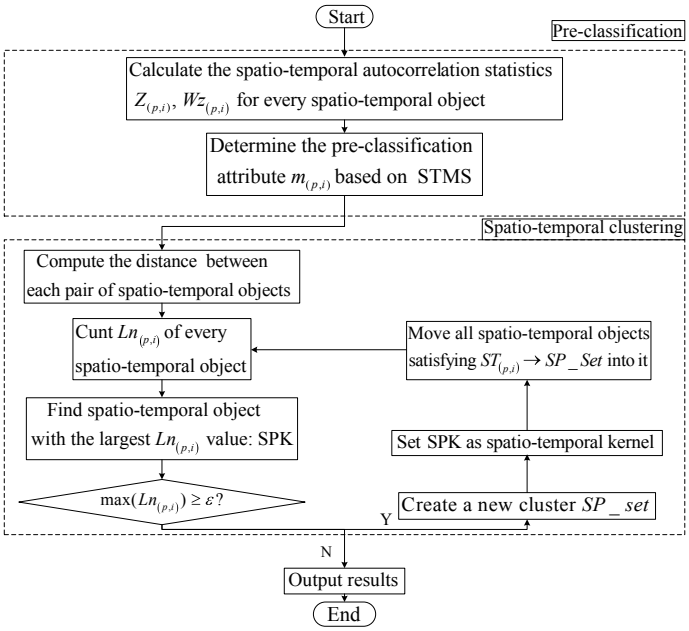


Fig. 2 Spatio-temporal clustering process combining pre-classification with STMS

(2) If $\max(Ln_{(p,i)})$ is less than ε, no candidate spatio-temporal kernel remains. So the clustering process is finished with all the spatio-temporal objects left considered to be discrete points.

3.2.2 Parameter Calibration

As a threshold to determine whether a spatio-temporal object is an candidate spatio-temporal kernel or not, with the increase of ε value, there are fewer candidate spatio-temporal kernels, which decreases the spatio-temporal clusters and increases the discrete points in clustering result. While in clustering analysis, it's usually wished that some kind of stable patterns with both few clusters and discrete points are found, so a series of exploratory clustering computations with different values of ε are conducted to find the proper value of it.

Assume that B is the collection of all possible values of ε. C and O represent the collections of clusters and discrete points numbers respectively through exploratory clustering computations. When ε taking value of B_i in B, the corresponding numbers of clusters and discrete points are C_i and O_i respectively, so the utility of this exploratory clustering with ε taking value of B_i can be calculated as Eq. (5).

$$U_{(B_i)} = \frac{C_i - Mean(C)}{Sd(C)} + \frac{O_i - Mean(O)}{Sd(O)} \tag{5}$$

where $Mean(C)$ and $Sd(C)$ are the average value and standard deviation of cluster numbers respectively. Similarly, $Mean(O)$ and $Sd(O)$ are average value and standard deviation of the numbers of discrete points. The utility of clustering result is hoped to be as small as possible, so there will be few clusters and discrete points.

4 Case Studies

Case studies of three urban expressways from the Second Ring Road to Fourth Ring Road in Beijing are carried out with the proposed spatio-temporal clustering method. The parameter ε is firstly calibrated through a series of exploratory clustering computations. And then, the clusters for 4 kind of traffic condition are presented and discussed.

4.1 Setup and Parameter Calibration

For each of the 3 expressways, two directions including the inner-ring (clockwise direction) and the outer-ring (counterclockwise direction) are studied respectively. The studied attribute is the average relative speed of Mondays in 2012. The studied expressways are all consist of road sections laid end to end, so every spatio-temporal object may have 8 spatio-temporal neighbors at least as shown in Fig. 3. Then the parameter ε can possibly take integer values from 1 to 8 to ensure that there is no spatio-temporal object certainly to become a single cluster or discrete point.

The results of exploratory clustering are presented in Table 2. For the average number of clusters or discrete points, there are no significant differences between

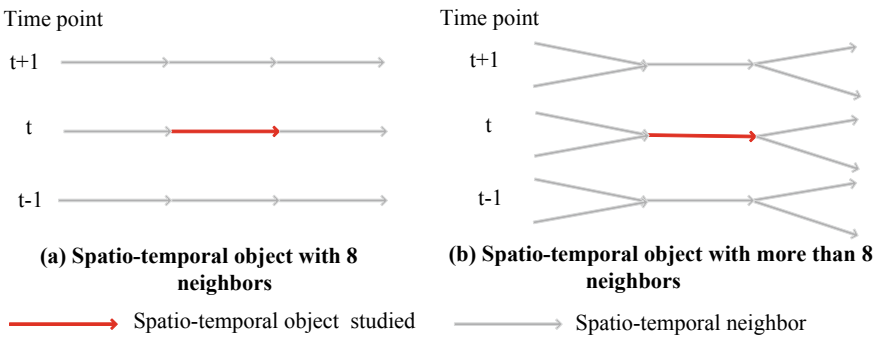


Fig. 3 Spatio-temporal object and its neighbors in space-time

Table 2 Results of exploratory clustering for the studied three urban expressways

Expressway	Direction	Average number of clusters	Average number of discrete points	ε value
Second ring road	Inner ring	16.63	148.25	3
	Outer ring	17.88	155.13	2
Third ring road	Inner ring	16.50	124.50	2
	Outer ring	19.75	172.88	2
Fourth ring road	Inner ring	23.00	249.00	2
	Outer ring	25.38	263.50	2

the inner-ring and outer-ring of each expressway. However, the average numbers of clusters and discrete points show a tendency of increase from urban center to surrounding suburbs (from the Second Ring Road to Fourth Ring Road). This is because that, the land utilization is usually more intensive in urban centers, so there is a higher spatio-temporal aggregation for same kinds of traffic condition in urban centers than the suburbs.

On the other hand, the proper ε value in Table 2 decreases with the location away from urban centers, which is contrary to the rule of the average number of clusters or discrete points. This is because that in suburbs with low spatio-temporal aggregation of traffic, a low ε value is enough to get the best clustering result, while a strict condition with high ε value can also get effective clustering result in urban centers where there are high spatio-temporal aggregation of traffic condition.

4.2 Clustering Results Combining Pre-classification

On the basis of the calibrated ε values for both directions of the 3 urban expressways, spatio-temporal clustering analyses are conducted. The spatio-temporal cloud maps of clustering results are shown in Fig. 4. According to STMS, traffic conditions are pre-classified into four types: homogenous uncongested traffic, heterogeneous uncongested traffic, homogenous congested traffic and heterogeneous congested traffic.

In Fig. 4, clusters of homogenous uncongested traffic are mainly distributed at the midnight between 0:00 and 6:00, while homogenous congested traffic clusters congregate in peak hours. The heterogeneous traffic clusters scatter around the rush hours in the daytime as some kind of transitional state between homogenous uncongested and congested traffic. While in space, with the location getting away from urban centers, the cluster size of homogenous congested traffic decreases, with the sizes of the clusters for discrete point and homogenous uncongested traffic increasing. This illustrates that, with the location away from urban centers with intensive land utilization, the spatio-temporal correlation of traffic condition gradually becomes weakened, and the traffic congestion getting alleviated too.

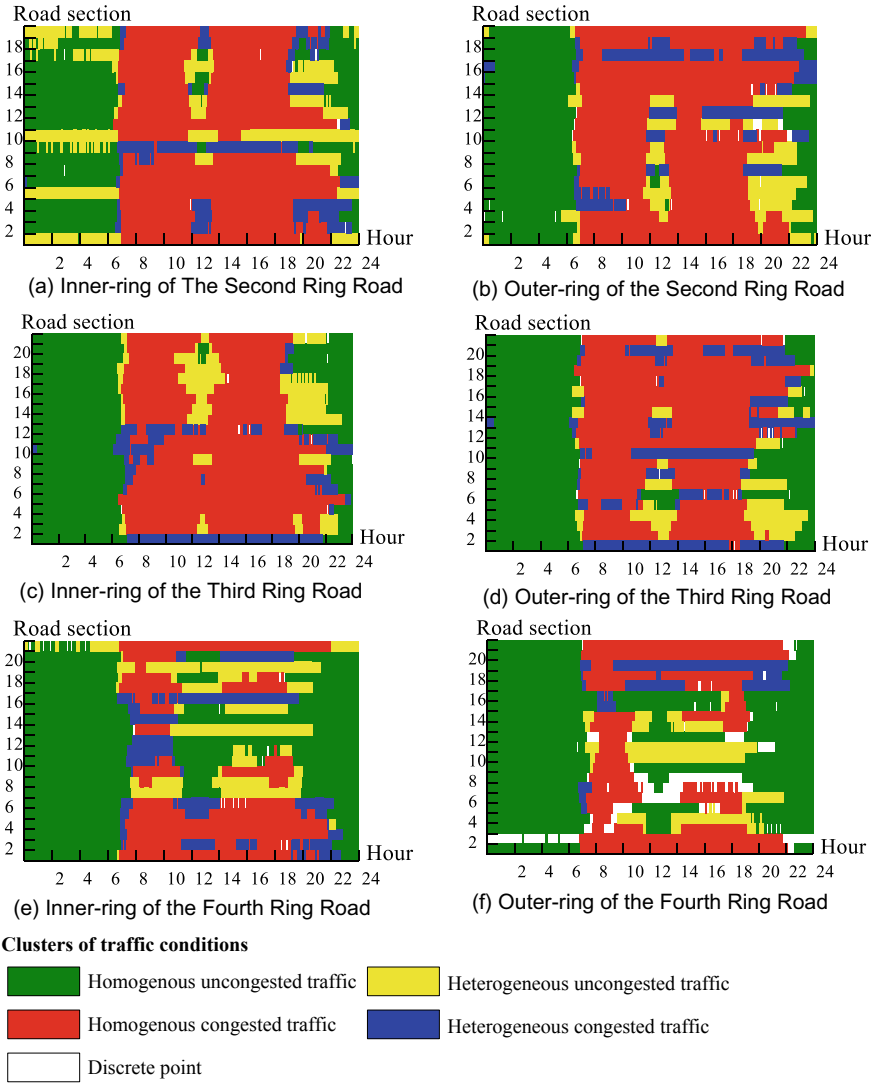


Fig. 4 Spatio-temporal cloud maps of traffic condition clusters for 3 expressways

Furthermore, road sections with heterogeneous uncongested traffic in Fig. 4 tend to become congested earlier and get uncongested later than others, which are bottlenecks in the road network. While road sections with heterogeneous congested traffic become congested later and get uncongested earlier, which play the role of dredging congestion during peak hours. Through a closer inspection, almost every road

section presents only one kind of heterogeneous traffic (the congested and uncongested) in the whole day, which illustrates that each road section in the network plays a relatively stable role in undertaking the daily traffic load.

4.3 Analysis and Discussion

In order to compare 4 kinds of traffic condition clusters with the realistic travel behaviors in the road network of Beijing, the moving trajectories of taxis with passenger on Mondays in November 2012 are also studied, as shown in Fig. 5. In detail, the trajectories of taxis in 4 durations: midnight, morning peak, noon and evening peak, are selected to study the daily change of travel behavior distribution.

In Fig. 5a, it can be found that travel behaviors are sparse between 0:00 and 2:00, so almost all road sections present homogenous uncongested traffic in Fig. 4. On the other hand, the aggregation of travel behaviors at peak hours in Fig. 5b and d

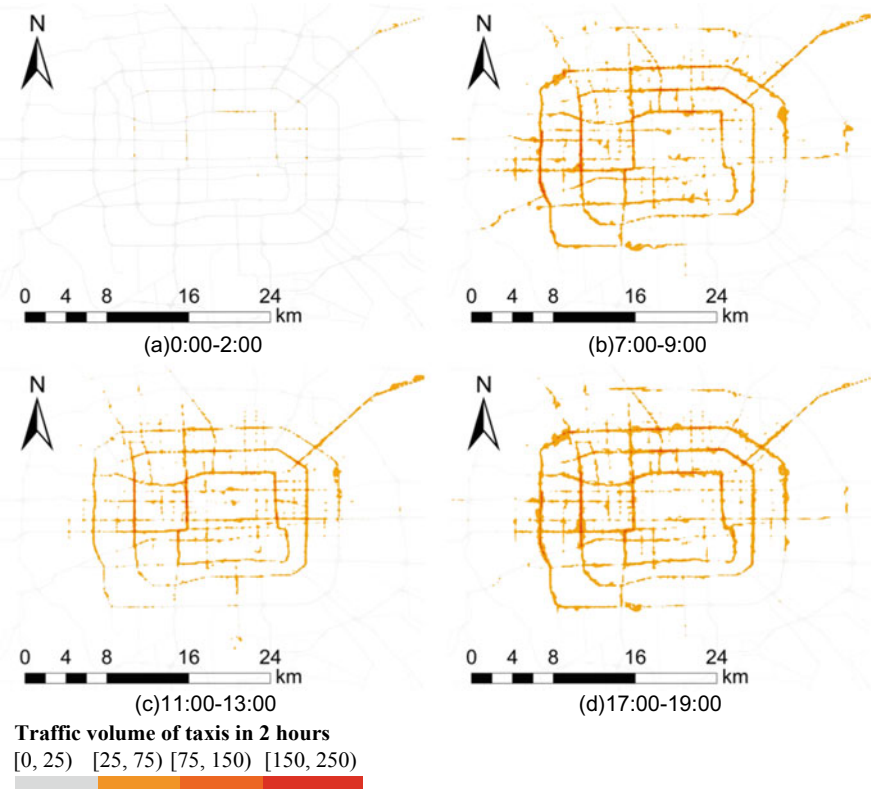


Fig. 5 Moving trajectories of taxis with passenger in the road network of Beijing

is stronger than other times, which causes that clusters of homogenous congested traffic are mainly distributed in the peak hours in Fig. 4. While between 11:00 and 13:00, travel behaviors become less congregated than peak hours, so the spatio-temporal correlation of traffic condition becomes weakened and traffic congestion gets alleviated.

While in space, according to the urban land use characteristics, there are more intensive points of traffic generation and attraction at urban centers than the suburbs, so a large amount of trajectory points congregate at the urban centers in Fig. 5b and d. Consequently, in urban centers, the correlation of traffic condition is higher than the suburbs. With the location away from urban centers, travel behaviors get sparser, so the spatio-temporal correlation of traffic condition becomes weakened gradually, and the traffic congestion gets alleviated as well in Fig. 4.

At the midnight (Fig. 5a), the volume of taxis in two hours is less than 25 at most road sections, but it may exceed 75 at some road sections in the east of the city. Further investigation reveals that, these road sections are near the famous Sanlitun Bar Street, where intensive travel behaviors last late into wee hours. This explains why road Sect. 5 (from Dongzhimenqiao to Dongsishitiaoqiao) and road Sect. 10 (from Guangqumenqiao to Guangmingqiao) at the east of the Second Ring Road in Fig. 4a are with heterogeneous uncongested traffic, whose traffic capacities should be improved to alleviate congestion.

5 Conclusions

Research on the traffic condition and its spatio-temporal distribution are crucial issues for Advanced Traffic Management Systems. A spatio-temporal clustering method combining pre-classification of traffic condition with STMS is proposed. Then the case studies of 3 urban expressways in Beijing are performed to verify the effectiveness of the proposed methodology. Through the case studies, some findings can be summarized as follows.

(1) With the location away from the urban centers, travel behaviors become more sparse, so traffic congestion get alleviated and the spatio-temporal correlation of traffic condition becomes weakened as well.
(2) When congestion forming and dissipating, travel behaviors become less congregated, so the traffic conditions of the whole network get heterogeneous in space-time. Road sections with heterogeneous uncongested traffic are bottlenecks of the network whose traffic capacities should be improved, while others with heterogeneous congested traffic can play congestion dredging role during peak hours.

So, the spatio-temporal clustering method combining STMS can not only effectively reveal the relation of traffic demand to road network facilities, but also recognize the road sections where congestion originates or gets alleviated in the network. This provides foundations for traffic managers to alleviate congestion and improve

urban transport services. Further studies are being considered to construct the quantitative model for the degree of traffic dredging or bottleneck effect of specific road section.

Acknowledgements The authors are grateful to the National Key R&D Program of China (2017YFB1200700). The authors also thank the anonymous reviewers and the editor for their suggestions to improve this paper.

References

1. BJTRC (Beijing Transportation Research Center) (2015) Beijing transport annual report (in Chinese)
2. Curry L (1970) Univariate spatial forecasting. Econ Geography 46:241–258
3. Cliff AD, Ord JK (1975) Space-time modelling with an application to regional forecasting. Trans Inst Br Geogr 64:119–128
4. Black WR (1992) Network autocorrelation in transportation network and flow systems. Geogr Anal 24(3):207–222
5. Chandra SR, Al-Deek H (2008) Cross-correlation analysis and multivariate prediction of spatial time series of freeway traffic speeds. Transp Res Rec 2061(9):64–76
6. Ma XL, Tao ZM, Wang YH et al (2015) Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp Res Part C 54:187–197
7. Dowling R, Skabardonis A, Carroll M, Wang Z (2004) Methodology for measuring recurrent and nonrecurrent traffic congestion. Transp Res Rec 1867:60–68
8. Varaiya P (2007) Finding and analyzing true effect of non-recurrent congestion on mobility and safety
9. Anbaroglu B, Heydecker B, Cheng T (2014) Spatio-temporal clustering for non-recurrent traffic congestion detection on urban road networks. Transp Res Part C 48:47–65
10. Kerner BS (2004) Three-phase traffic theory and highway capacity. Phys A 333:379–440
11. Kerner BS (2004) The physics of traffic: empirical freeway patten features, engineering applications and theory. Springer
12. Salazar E, Dunson DB, Carin L (2013) Analysis of space-time relational data with application to legislative voting. Comput Stat Data Anal 68:141–154
13. López-hernández FA, Chasco C (2007) Time-trend in spatial dependence: specification strategy in the first-order spatial autoregressive model. Estudios de Economia Aplicada 25(2):631–650
14. Hardisty F, Klippel A (2010) Analysing spatio-temporal autocorrelation with lista-viz. Geogr Inf Sci 24(10):1515–1526
15. Chen SK, Wei W, Mao BH, Guan W (2013) Analysis on urban traffic status based on improved spatio-temporal Moran's I. Acta Phys Sin 62(14):148901
16. Pfeifer PE, Deutsch JA (1980) Three-stage iterative procedure for space-time modeling. Technometrics 22(1):35–47
17. Kamarianakis Y, Prastacos P (2005) Space-time modeling of traffic flow. Comput Geosci 31(2):119–133
18. Deng M, Liu QL, Wang JQ, Shi Y (2012) A general method of spatio-temporal clustering analysis. Scientia Sinica (Informationis) 42(1):111–124
19. Kulldorff M, Heffernan R, Hartman J et al (2005) A space-time permutation scan statistics for disease outbreak detection. PLoS Med 2(3):216–224
20. Gaudart J, Poudiougou B, Dicko A et al (2006) Space-time clustering of childhood malaria at the household level: a dynamic cohort in a mali village. BMC Public Health 6:1–13
21. Wang M, Wang AP, Li AB (2006) Mining spatial-temporal clusters from geo-database. Lecture notes in artificial intelligence, vol 4093, pp 263-270

22. Pei T, Zhou CH, Zhu AX et al (2010) Windowed nearest neighbour method for mining spatio-temporal clusters in the presence of noise. Int J Geogr Inf Sci 24(6):925–948
23. Kulldorff M, Hialmars U (1999) The Knox method and other tests for space-time interaction. Biometrics 55(2):544–552
24. Zaliapin I, Gabrielov A, Keilis-borok V et al (2008) Clustering analysis of seismicity and aftershock identification. Phys Rev Lett 101(1):018501
25. Steinbach M, Tan PN, Kumar V et al (2001) Clustering earth science data: goals, issues and results. In: Kamath C (ed) Proceedings of the 4th KDD workshop on mining scientific datasets in conjunction with 7th ACM SIGKDD international conference on knowledge and data mining. ACM Press, San Francisco, pp 1–8
26. Birant D, Kut A (2007) ST-DBSCAN: an algorithm for clustering spatial-temporal data. Data Knowl Eng 60(1):208–221
27. Moran PAP (1948) The interpretation of statistical maps. Biometrika 35:255–260
28. Anselin L (1988) Spatial econometrics: method and models. Kluwer Academic Publishers, Dordrecht
29. Cliff AD, Ord JK (1973) Spatial autocorrelation. Pion Limited, London
30. Cliff AD, Ord JK (1981) Spatial processes: models and applications. Pion Limited, London
31. Griffith D (2003) Spatial autocorrelation and spatial filtering. Springer, Berlin
32. Anselin L (1995) The local indicators of spatial association-LISA. Geogr Anal 27:93–115

Passenger Flow Prediction for Urban Rail Transit Stations Considering Weather Conditions

Kangkang He, Gang Ren and Shuichao Zhang

Abstract Precise prediction of urban rail transit passenger flow is essential for the development of organizing plans by the rail transit management and operation department, and also is the fundament to achieving passenger transport guarantees. This study collected Ningbo rail transit Route 2 passenger flow data and candidates of key driving factors including station type, population and employment position density, transfer facilities, main land area within an 800 m radius, particularly considering weather conditions, and then Random Forest was applied for passenger flow prediction. The prediction results show that the models considering the weather factors is superior to the models without consideration, mean absolute deviation (MAD) and mean absolute percentage deviation (MAPD) are reduced by 14.40 and 57.55%, respectively. The model involved weather factors is more accurate under hot and heavy rain weather conditions. Employment position, population density and commercial service facilities land area within an 800 m radius of the station, are the most important factors influencing the passenger flow, while average temperature is more likely to affect the passenger flow than precipitation. These results suggest that the passenger flow forecasting model based on random forest can achieve rapid calculation under different weather conditions, and provide important data basis for urban rail transit passenger flow density warning, passenger flow guidance and operation scheduling.

Keywords Rail transit · Passenger flow prediction · Random forest · Weather condition

K. He · G. Ren (✉)
School of Transportation, Southeast University, Nanjing 210096, China
e-mail: rengang@seu.edu.cn

S. Zhang
School of Architecture and Transportion Engineering, Ningbo University of Technology, Ningbo 315211, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_51

1 Introduction

Urban rail transit has become one of the most important transport modes for commuter passengers in the city due to its advantages of large capacity, low energy consumption, safety, punctuality and high-speed [1]. The rail transit station serves as a platform for direct interaction between the rail transit system and gathering passengers. Therefore, accurate prediction of station passenger flow is of great significance in improving the operational safety and efficiency of rail transit stations, and thus guiding strategies for enhancing the transport capacity of stations during peak hours.

The regularity analysis and forecasting method of rail transit passenger flow has become a research hotspot in recent years. Zhang et al. [2] proposed a short-term passenger flow forecasting model based on a modified Kalman filtering, which could improve the real-time performance of forecasting information, reduce the average absolute error, and further improve the forecasting accuracy compared with the traditional Kalman filtering. A recent study applying moving average method (MAM) based on passenger flow data in Shanghai Route 1 (China) reported that the prediction accuracy of MAM was higher than those derived from support vector machine, back propagation neural network, wavelet neural network and wavelet combination support vector machine [3]. In addition, Bayesian network was also frequently applied, with which Yuan et al. [4] found that short-term flow in a certain station was affected by the passenger flow at adjacent stations and adjacent time phases.

In passenger flow prediction model establishment, it is crucial to select key driving factors for ensuring the accuracy. Many scholars have established models incorporating multiple factors, and believe that changes in passenger flow at rail transit stations are mainly affected by population and employment position density [9], land use [5], transfer facilities [10], station type [4], price [6], service level [7]. Taylor et al. [8] found that passenger flow were determined by both intrinsic (such as price and service level) and extrinsic (such as land use and transfer facilities) factors, suggesting that intrinsic factors are as important as intrinsic factors. Later, a study considering more potential factors demonstrated that population density, median household income, employment position density, accessibility by walking, parking fee, local vehicle travel time, destination area size, and center area position determined the passenger flow in Broward County, Florida (USA) [9]. Jun et al. [10] found that the population, employment position density, diversity of land use, and multimodal public transport convergence represented by the number of buses exhibited a positive effect on the passenger flow in the core area or pedestrian attraction area of Seoul (South Korea) rail transit stations.

Occurrences of unfavorable weather conditions (wind, rain, snow, ice, etc.) can affect residents' travel choices including route, mode, and timing of travel, thus leading to changes in the passenger flow of the rail transit system [11–14]. Guo et al. [15] investigated the passenger flow of bus and rail transits in Chicago (USA) with linear regression model considering temperature, wind strength, and occurrences of rain and fog, suggesting that adverse weather conditions lead to a lower passenger flow compared with that under regularity over the week, particularly in the working

days. The impact of weather on the passenger flow was reported to vary with public transport mode (rail transit, bus), timing (weekdays and weekends) and season in New York (USA), but rainy weather generally showed a negative impact on various modes of transit regardless of timing and season [16]. However, in the case of weak rain, passenger flow was not substantially influenced, and sensitivity of passenger flow gradually rose with the incensement of rain [17]. It seems that unfavorable weather conditions negatively influence the passenger flow and the influencing extent of different weather factors is likely distinct.

The primary objective of this study is to predict the passenger flow of urban rail transit stations considering weather factors. More specifically, we aims: (1) to develop the Random Forest model considering the weather factors and the model without consideration; (2) evaluate if predicions from weather-considered model are more accurate from the model without consideration; and (3) evaluate how various factors affect the accuracy of the prediction model. Findings of this study can be directly used by rail transit authorities to accurately predict the rail transit stations passenger flow and further guide establishment of organizing plans, passenger flow warning, and emergency passenger flow organization under different weather conditions.

2 Data Collection

The study area is located in Ningbo City, a middle city in Zhejiang Province, which is one of the most economically developed provinces in China (Fig. 1). The area and population of Ningbo city are 1033 km^2 and 7.64 million, respectively. In the study area, multi-mode public transport system including rail transit, customized shuttle bus, and public bicycle has been extensively developed. Moreover, the total

Fig. 1 Study area location and sampling points

number of motor vehicles has rapidly increased to 2.5 million, and heavy traffic exists throughout the city.

The passenger flow data (daily) of Route 2 of Ningbo Rail Transit is provided by Ningbo Rail Transit Operation Co., Ltd. Route 2 was started construction on December 23, 2010 and began operation on September 26, 2015. The route with a total length of 28.35 km and a total of 22 stations soon became the backbone of the southwest-northeast direction of Ningbo City. The average distance between neighboring stations is 1.33 km. According to the needs of the study, data on the passenger flow of 22 stations (Route 2) from January 1st to November 30th in 2017 and built environment of rail trasit stations were collected (Fig. 2).

Selecting key factors covering as many terms as possible can effectively improve the performance of prediction model. Based on a review of the literatures [5–10, 18–20], we selected a priori 16 variables (as listed in Table 1) to explain the spatial variability in the passenger flow.

3 Methodology

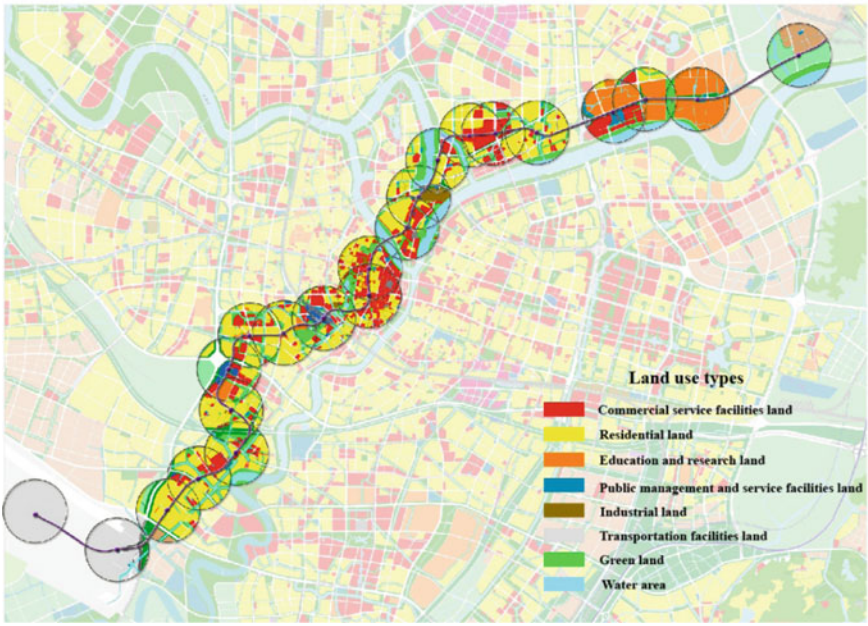
3.1 Random Forest

Random Forest (RF) studied in this paper is a combined classifier based on classification and regression tree proposed by Breiman [21]. Its superior performance has made it widely used in many fields such as biology [22], medicine [23], economics [24], and management [25]. In recent years, RF has become more and more widely used in the transportation field, and has been used in fuel consumption forecasting [26], traffic incident detection [27], transportation mode choice [28], etc.

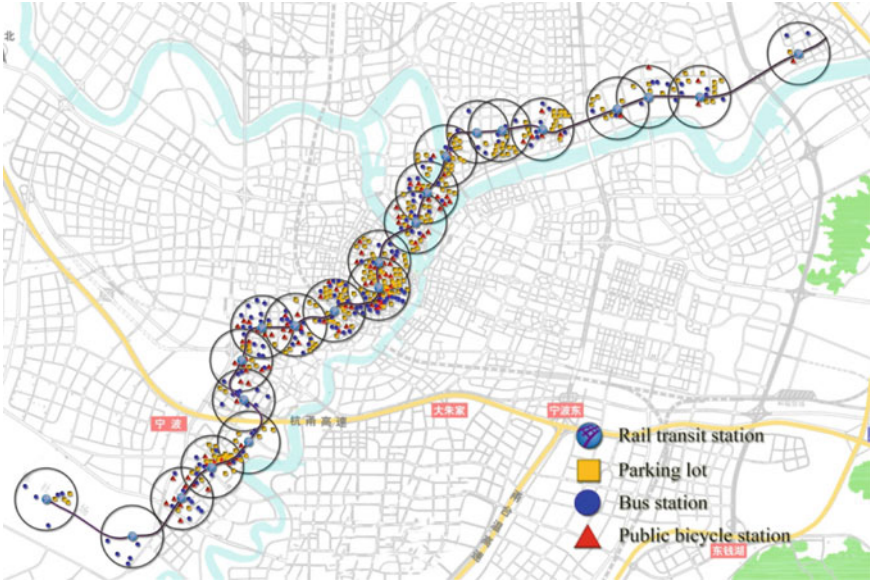
RF is a combined classifier which is implemented in the following three steps [21]:

- The bootstrap sampling technique is used to extract n training sets from the original data set. The size of each training set is about two-thirds of the original data set.
- A Classification and Regression Tree is established for each bootstrap training set. A total of ntree of decision trees are constructed to form a forest. These decision trees are not pruned. In the growth process of each tree, instead of selecting the optimal attribute among all M attributes as an internal node to split, an optimal attribute is selected from the randomly selected mtry $\leq$M attributes for branching.
- The prediction results of the decision trees of the ntree are collected, and voting is used to determine the categories of the new samples.

In accordance with the above steps, the RF algorithm is different from the decision tree, which introduces two random sources during the construction of each tree: the first random source is a new training set generated randomly from the original training set, followed by each tree growing on this randomly generated training set and without pruned; the second random source is a randomly entered variable when



(a) Land use types within 800 meters radius of rail transit stations.



(b) Transfer facilities within 800 meters radius of rail transit stations.

Fig. 2 Land use types and transfer facilities within an 800 m radius of rail transit stations. **a** Land use types within an 800 m radius of rail transit stations. **b** Transfer facilities within an 800 m radius of rail transit stations

Table 1 Variables selected for model calibration

	Unit	Type	Max.	Min.	Mean
Station type					
Transfer station	–	Binary			
Terminal	–	Binary	–	–	–
Regular station	–	Binary	–	–	–
Population and employment position density					
Population density	Person/km^2	Continuous	26,000	1100	8585.8
Employment position density	Employment position/km^2	Continuous	23,000	500	6619.5
Transfer facilities within an 800 m radius of rail transit stations					
Number of bus routes	–	Continuous	63.00	2.00	16.63
Number of public bicycle stations	–	Continuous	15.00	0.00	4.60
Number of parking lots	–	Continuous	20.00	0.00	5.74
Main land area within an 800 m radius					
Residential land area	10,000 m^2	Continuous	142.81	0.00	82.69
Commercial service facilities land area	10,000 m^2	Continuous	194.10	0.00	45.06
Industrial land area	10,000 m^2	Continuous	26.07	0.00	1.91
Public management and service facilities land area	10,000 m^2	Continuous	94.44	6.26	26.39
Weather conditions					
Average temperature	°C	Continuous	34.00	2.00	19.96
Precipitation	mL/day	Continuous	30.60	0.00	5.27
Working day/holiday					
Working day	–	Binary	–	–	–
Holiday	–	Binary	–	–	–

the node splits. These two random sources make the RF algorithm error rate more stable, which overcome the deficiencies of a single decision tree and reflect the advantages of the integration of the RF algorithm.

We used the "randomForest" package in the R environment to create two RF models based on the 7348-sample training dataset [29]. The model considering weather condition involved 16 factors as listing in Table 1, while the model without weather condition has 14 factors removing average temperature and precipitation. Three main parameters: the number of decision trees in the forest (ntree), the number of randomly

selected attributes of internal nodes (mtry), and the minimum sample size of the final node (nodesize), are set to 1000, 16, and 5, respectively [30].

3.2 Measures of Goodness of Fitting

A 10-fold cross-validation was performed to tune and evaluate the models, i.e., the input samples (7348 sets of data) were randomly partitioned into 10 subsets and models are built on 9 subsets and validated using the rest subset. This cross-validation process is then repeated 10 times, with each of the 10 subsets used exactly once as the validation data. The accuracy of models is measured as root mean squared error (RMSE), calculated as the standard deviation of the differences between the predicted and observed values, representing the average prediction error in the same unit of original data (in this case, days). The most parsimonious model (in order to avoid overfitting to the input training data) within one standard deviation of the optimal accuracy were chosen as the final model [31].

In order to evaluate the prediction effect, mean absolute deviation (MAD), mean square error (MSE), mean absolute percentage deviation (MAPD), mean square percentage error (MSPE), and equal coefficient (EC) are introduced to evaluate the forecasted value [2]. Let the observation data be yi and the prediction value be $\hat{y}i$, the measures are formulated as follows:

$$\mathrm{MAD} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}i - yi\right| \tag{1}$$

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}\left(\hat{y}i - yi\right)^2}{n} \tag{2}$$

$$\mathrm{MAPD} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}i - yi}{yi}\right| \tag{3}$$

$$\mathrm{MSPE} = \frac{\sum_{i=1}^{n}\left(\frac{\hat{y}i - yi}{yi}\right)^2}{n} \tag{4}$$

$$\mathrm{EC} = 1 - \frac{\sqrt{\sum_{i=1}^{n}\left(\hat{y}i - yi\right)^2}}{\sqrt{\sum_{i=1}^{n} yi^2} + \sqrt{\sum_{i=1}^{n}\hat{y}i^2}} \tag{5}$$

The MAD and MAPD provide a measure of the average misprediction of the model. The MSE and MSPE are typically used to assess the error associated with

a prediction. A smaller value of MAD, MAPD, MSE or MSPE suggests that, on average, the model predicts the observed data better. EC is the fitting degree between predicted value and observed value. When EC $\geq$ 0.9, it shows that it is a good fitting.

4 Results and Discussion

4.1 Predicting Accuracy Comparison

The prediction performances of the two RF models which one is considered weather condition and the other is not consideration were compared. The passenger flow was predicted for each station given the data from the aforementioned test set. The measures of prediction accuracy introduced in Sect. 3.2 were computed for the two models. MAD, MSE, MAPD, MSPE and EC for the model involved weather factors are less than those for the model without weather factors. The MAD, MSE, MAPD, MSPE and EC for the weather model were reduced by 14.40%, 2.09%, 57.55%, 316.95% and 6.29%, respectively when compared to the no weather factor model. The results indicate that the model with weather factors produces more accurate predictions for passenger flow (Fig. 3; Table 2).

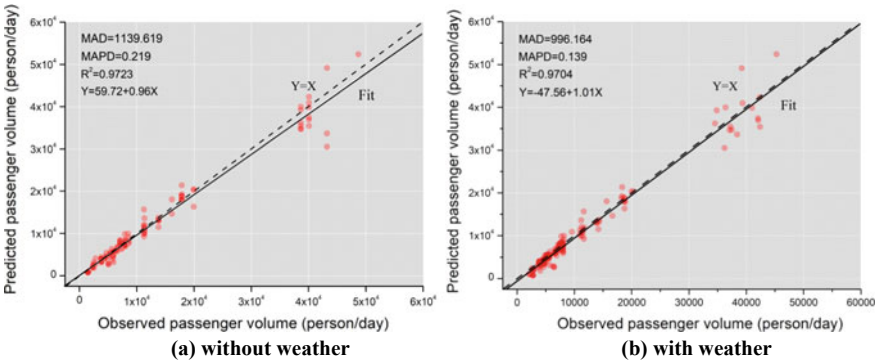


Fig. 3 Performances of two RF models. **a** Without weather **b** With weather

Table 2 Comparison of predicting accuracy between models

	MAD	MSE	MAPD	MSPE	EC
Without weather	1139.6	3.38E6	0.219	0.246	0.879
With weather	996.2	3.32E6	0.139	0.059	0.938

4.2 Model Performance Under Different Weather Conditions

We considered four types of weather conditions to figure out the fitting goodness of two models. Four weather conditions were defined as follows: hot weather: the highest temperature exceeds 35 °C, cold weather: the lowest temperature is lower than 0 °C, heavy rain: the daily precipitation reaches 20 ml (medium rain level), pleasant weather: the highest temperature and lowest temperature are both between 10–30 °C, no rain [15].

The prediction performances of the two RF models were compared. Under hot and heavily rainy weather conditions, the model involving weather factors is obviously better than the non-weather model, while two models performance similar under cold and pleasant weather. Two models both perform worst under cold weather, and predicting accuracy measure indices such as MAD are 4.18 and 7.33% lower compared to pleasant weather, respectively (Fig. 4; Table 3).

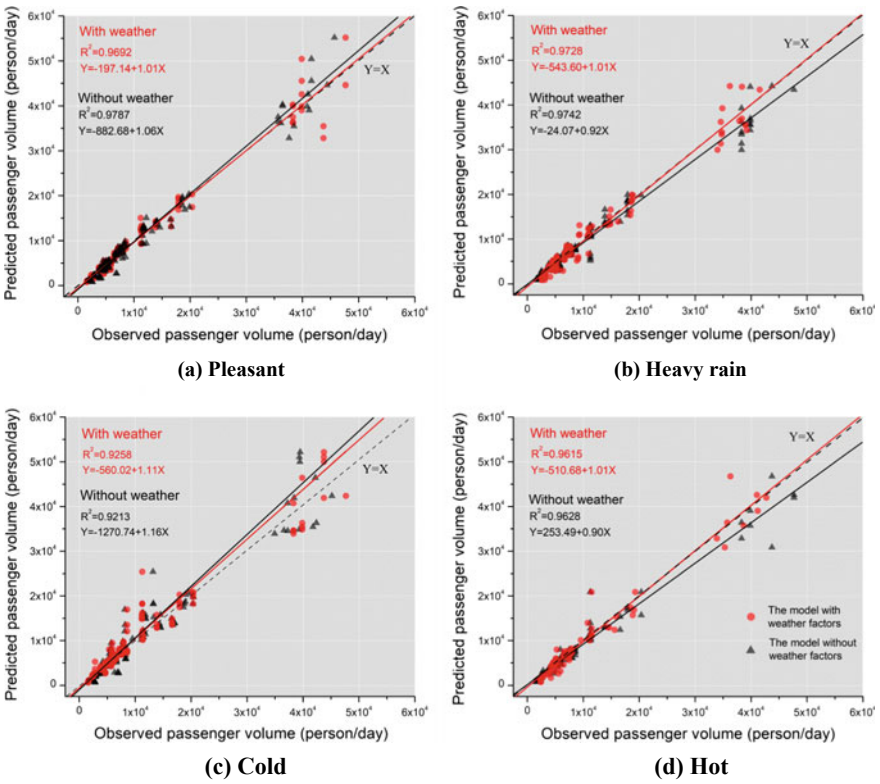


Fig. 4 Performances of two RF models under different weather conditions. **a** Pleasant **b** Heavy rain **c** Cold **d** Hot

Table 3 Comparison of predicting accuracy between models under different weather conditions

	MAD	MSE	MAPD	MSPE	EC
Pleasant					
Without weather	1129.7	3.64 E6	0.238	0.253	0.868
With weather	975.4	3.45 E6	0.156	0.067	0.901
Hot					
Without weather	1133.5	3.62 E6	0.234	0.212	0.870
With weather	921.9	2.94 E6	0.129	0.058	0.941
Cold					
Without weather	1176.9	5.43 E6	0.217	0.231	0.870
With weather	1046.9	4.03 E6	0.173	0.096	0.918
Heavy rain					
Without weather	1173.0	3.46 E6	0.245	0.214	0.845
With weather	912.7	2.71 E6	0.133	0.057	0.928

4.3 Relative Importance of the Predictors

In order to evaluate the relative importance of each predictor in the models, we established 16 new models by removing one single factor at each time. Changes in prediction accuracy of 16 new models compared to the original model including all factors was analyzed to reveal relative contribution of each predictor to the classification model, thus disentangling the strength of the predictors (Table 4).

Table 4 The relative importance of 16 predictors

Missing predictor	The average reduction in accuracy (%)	Missing predictor	The average reduction in accuracy (%)
Employment position density	19.96	Industrial land area	17.10
Population density	19.50	Average temperature	16.58
Commercial service facilities land area	19.44	Transfer station	16.44
Public management and service facilities land area	19.04	Regular station	16.38
Residential land area	18.87	Holiday	15.25
Number of parking lots	18.70	Working day	15.03
Number of bus routes	17.72	Precipitation	13.67
Number of public bicycle stations	17.54	Terminal	12.41

After removing employment position density, the prediction accuracy decreased the most, reaching 19.96%, and thus the position density is the most important factor determining the passenger flow. The second and third factors are the population density and commercial service facilities land area within 800 m radius of rail transit stations, reaching 19.50 and 19.44%, respectively. Previous literatures [9, 10] also support this result. The factors ranked in the fourth to ninth places are commercial service facilities land area, public management and service facilities land area, residential land area, the number of public parking lots, the number of bus lines, the number of public bicycle stations, industrial land area. It seems average temperature is more likely to affect passenger flow than precipitation. This result is consistent with previous literatures [17, 32]. The fact is that high or low temperature would change trip decisions. In addition, hot and cold weather conditions overlap with the students winter and summer vacation, which has a additive effect on passenger flow. Sprinkle only increases travel time and passengers change their travel start time rather than change transportation mode. However, heavy rain will lead to a sudden change in the passenger flow. Literature [9] found out the passenger flow of rail transit increased in Nanjing (China), for passengers diverting from other transport modes. However the passenger flow of Chicago (USA) rail transit system reduced, because passengers diverting to private cars or decide not go out [15]. The influence degree of terminal is the smallest, mainly due to the fact that the passenger flow of Ningbo rail transit terminals are not significantly different from other stations, therefore it is difficult to find out the pattern.

5 Conclusion

Accurate prediction of urban rail transit passenger flow is of great significance to disperse the passenger flow and rationally allocate operational resources. The RF modeling results showed that RF prediction model considering the weather factor is superior to the model that does not consider, and the model involved weather factor is more accurate in predicting the hot and heavy rain weather. In addition, the results demonstrated the RF passenger flow forecasting model can achieve rapid calculation considered different weather conditions, and provide important data basis for urban rail transit passenger flow density warning, passenger flow guidance and operation scheduling.

Some results were limited by the data that was available for analysis. For example, we only predicted the daily passenger flow for rail transit stations and did not analyse short-time (such as 5 min or 15 min) passenger flow which is more meaningful to improve the operational security and efficiency of stations during peak hours. In addition, spatial-temporal heterogeneity was not included in the analysis and did not find out the impact of variables on different stations at different time phases. The authors recommend that future studies may focus on collecting more data to enrich the results and analyses on the study topic.

Acknowledgements This research was funded by the National Natural Science Foundations of China (51308311, 51578149) and the Philosophy and Social Science Program of Zhejiang Province, China (18NDJC107YB). The authors wish to thank students from the Ningbo University of Technology in P. R. China, who were involved in the data collection process. And special thanks to Ningbo Rail Transit Operation Co., Ltd. provided operation data.

References

1. Song L, Li Q, List G et al (2017) Using an AHP-ISM based method to study the vulnerability factors of urban rail transit system. Sustainability 9(6):1065
2. Zhang Z, Zhang D, Jia J et al (2017) Short-term passenger flow forecasting of rail transit platform based on improved Kalman filter. J Wuhan Univ Technol (Transp Sci Eng) 6:974–977 (in Chinese)
3. Meng P, Li X, Jia H et al (2018) Short-time rail transit passenger flow real-time prediction based on moving average. J Jilin Univ (Eng Technol Ed) 48(2):448–453 (in Chinese)
4. Yuan J, Wang P, Wang Y et al (2017) A passenger volume prediction method based on temporal and spatial characteristics for urban rail transit. J Beijing Jiaotong Univ 41(6):42–48 (in Chinese)
5. Liu J, Zhang Y (2004) Analysis to passenger volume effect of land use along urban rail transit. Urban Transport of China
6. Jansson J (2000) Double dividend of efficient pricing of railway passenger transport. In: Asia Pacific conference and exhibition on transportation and the environment april Beijing
7. Barlach Y, Shiftan Y, Sheffer D (2007) Passengers attitude toward bus and rail: investigation of corridor with a similar level of service. In: 11th world conference on transport research
8. Taylor BD, Miller D, Iseki H et al (2009) Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. Transp Res Part A Policy Pract 43(1):60–77
9. Thompson G, Brown J, Bhattacharya T (2012) What really matters for increasing transit ridership: understanding the determinants of transit ridership demand in Broward county. Florida Urban Stud 49(15):3327–3345
10. Jun MJ, Choi K, Jeong JE et al (2015) Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. J Transp Geogr 48:30–40
11. Singhal A, Kamga C, Yazici A (2014) Impact of weather on urban transit ridership. Transp Res Part A Policy Pract 69(69):379–391
12. Kalkstein AJ, Kuby M, Gerrity D et al (2009) An analysis of air mass affects on rail ridership in three US cities. J Transp Geogr 17(3):198–207
13. Keay K, Simmonds I (2005) The association of rainfall and other weather variables with road traffic volume in Melbourne. Aust Accid Anal Prev 37(1):109–124
14. Brijs T, Karlis D, Wets G (2008) Studying the effect of weather conditions on daily crash counts using a discrete time-series model. Accid Anal Prev 40(3):1180–1190
15. Guo Z, Wilson NHM, Rahbee A (2008) Impact of weather on transit ridership in Chicago, Illinois. Transp Res Rec J Transp Res Board 2034(2034):3–10
16. Cravo VS, Cohen J, Williams A (2009) Impact of weather on transit revenue in New York City. Transp Res Board 88th Annu Meet
17. Lou S (2016) Study on the influence of rainfall on the passenger flow in station of urban rail transit. Southeast University
18. Zhao P, Lü B, Roo GD (2011) Impact of the jobs-housing balance on urban commuting in Beijing in the transformation era. J Transp Geogr 19(1):59–69
19. Peng ZR (2014) The jobs-housing balance and urban commuting. Urban Stud 34(8):1215–1235
20. Taylor BD, Fink C (2013) Explaining transit ridership: what has the evidence shown? Transp Lett 5(1):15–26

21. Breiman L (2001) Random forests. Mach Learn 45:5–32
22. Boulesteix A, Janitza S, Kruppa J et al (2012) Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wiley Interdisc Rev Data Min Knowl Discovery 2(6):493–507
23. Guo R, Wang Y, Yan H et al (2015) Analysis and recognition of traditional Chinese medicine pulse based on the hilbert-huang transform and random forest in patients with coronary heart disease. Evid Based Complement Altern Med 2015 (2015-6-9), 2015, 2015(6216, supplement):1–8
24. Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J (2007) Classification in ecology. Ecology 88:2783–2792
25. Uriarte RB, Tiezzi F, Tsaftaris SA (2016) Supporting autonomic management of clouds: service clustering with random forest. IEEE Trans Netw Serv Manage 13(3):595–607
26. Wang SL, Xie F, Zhu HY et al (2017) Research of aircraft fuel consumption based on random forest. Math Pract Theor
27. Liu Q, Lu J, Chen S (2013) Traffic incident detection using random forest. Transp Res Board 92nd Annu Meet
28. Sekhar CR, Minal M, Madhu E (2016) Multimodal choice modeling using random forest decision trees. Int J Traffic Transp Eng 6(3):356–367
29. Breiman L, Cutler A (2013) Breiman and cutler's random forests for classification and regression. R package version 46–7. Available: https://cran.r-project.org/web/packages/randomForest/
30. Qiu L, Kai W, Long W et al (2016) A comparative assessment of the influences of human impacts on soil Cd concentrations based on stepwise linear regression, classification and regression tree, and random forest models. PloS One 11(3)
31. Metcalf JL, Xu ZZ, Weiss S et al (2016) Microbial community assembly and metabolic function during mammalian corpse decomposition. Science 351(6269):158–162
32. Sumalee A, Uchida K, Lam WHK (2011) Stochastic multi-modal transport network under demand uncertainties and adverse weather condition. Transp Res Part C Emerg Technol 19(2):338–350

An Algorithm for Searching Freeway Speeding Unlicensed Vehicles

Lu-li Liang, Wen-hong Lv, Peng-fei Wang, Jia-li Ge and Guo-juan Wang

Abstract A timeless retrospective algorithm for finding unlicensed vehicles on highway speeding is proposed. The algorithm mainly uses the images captured by the freeway capture device to obtain the speed information of the vehicles, and the distance and speed of the vehicles. The time backtracking method is used to determine when the vehicle enters the entrance, then **screen the image to determine the illegal vehicle information.** The application of the algorithm is implemented through a C language program. The results show that freeway speeding unlicensed vehicle search algorithm can search for speeding unlicensed vehicles and alert freeways for unlicensed speeding.

Keywords Overspeeding · Unlicensed vehicle · Image recognition · Time backtracking

1 Introduction

With the increase of speed, the reaction time of driver and vehicle will be prolonged. In this case, the vehicle cannot be stopped in time. Overspeeding will stimulate other drivers, and they are prone to misjudgment in case of any emergencies. As a result, they may lose control of the vehicle, which interferes with the normal traffic. The greater the speed is, the more severe the consequences will be if there is any emergencies.

The freeway overspeeding vehicles to be searched are divided into two types, namely licensed vehicles and unlicensed vehicles. For the licensed vehicles, the vehicle information is mainly determined by vehicle recognition. In 2004, "General Technical Conditions of Intelligent Detection and Recording System for Road Vehicles" for the first time proposed the concept of Motoring and Controlling Vehicle, namely comparing and checking the vehicle recognition result of the motor license

L. Liang · W. Lv (✉) · P. Wang · J. Ge · G. Wang
College of Transportation, Shandong University of Science and Technology, Qingdao 266590, China
e-mail: zklwh@sdust.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_52

plate with the set vehicle number. It is one of the functions of motor vehicle monitoring [1]. To date, in-depth research on license plate recognition has been carried out [2], and mature products have been applied to the actual traffic management. For unlicensed vehicles, the recognition of brand, model, color can be made by comparing and checking the information of vehicles in violation of traffic rules to confirm the vehicle information. At present, the relatively mature vehicle recognition methods include loop coil detection [3], infrared detection [4], ultrasonic/microwave detection [5], etc. Nonetheless, they are easy to be affected by the weather and difficult to maintain. The image-based vehicle recognition technology has been gradually maturing. In 2015, Peng Bo et al. came up with a vehicle recognition method based on in-depth learning, which is accurate and stable with the advantages of independent learning of auto logo feature and direct input of image [6]. In 2016, Deng Liu et al. in view of the problem of vehicle recognition in freeway environment, designed the feature extraction algorithm in accordance with the Theory of Convolutional Neural Network. Meanwhile, the recognition system was built for vehicle recognition combined with the SVM classifier, which made improved in the recognition accuracy and speed [7]. In 2017, Song Huansheng et al. applied the in-depth learning objective classification algorithm to transform the target detection problem into objective dichotomy problem for the vehicles in the actual traffic scene, so as to achieve the detective recognition of vehicle targets [8]. In the premise of the vehicle recognition based on deep learning, the paper proposes an algorithm to find the unlicensed vehicle by determining the period in which the vehicle enters the freeway entrance.

2 Speeding Unlicensed Vehicle Search Algorithm

2.1 Principles of the Algorithm

The speed, location, image and other information of the overspeeding unlicensed vehicle are acquired directly by the freeway speed-measuring snapshot device. The information obtained is utilized to calculate the travel time of the vehicle, and the timing of the vehicle entering the freeway entrance is calculated, thus obtaining the time period when the overspeeding unlicensed vehicle enters the freeway entrance. The images taken at the freeway entrance during the time period were recognized by vehicle model recognition and marker recognition, and the information of vehicles in violation of traffic rules is determined (Fig. 1).

2.2 Unlicensed Vehicle Speed Information

To date, the freeway speed measurement is an organic combination of speed measurement technology and image acquisition technology. Through the real-time automatic

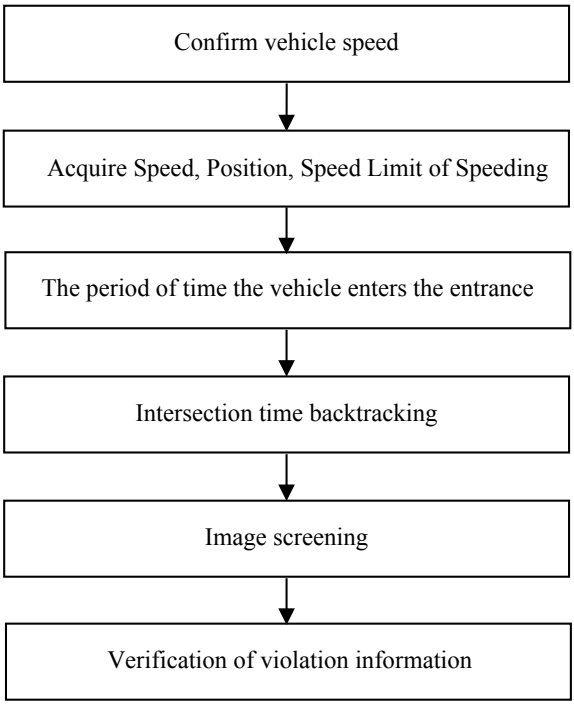


Fig. 1 Algorithm flowchart

measurement of the driving speed on the road and image shooting for overspeeding vehicles, the speed, time, location, speed limit and other information of the vehicles are recorded automatically, as is shown in Fig. 2. The acquisition of speed, position and speed limit of the algorithm proposed in the paper is achieved directly by the speed-measuring snapshot device of the freeway.

Fig. 2 Freeway speed capture device shooting pictures

2.3 Determination of Vehicle Travel Time

2.3.1 Determination of the Time for Ordinary Speed Measurement

The vehicle information such as speed and position can be obtained directly by means of freeway speed-measuring snapshot device. The time of the algorithm proposed in the paper can be calculated with the formula T = S/V, obtaining the time it takes for the vehicle to travel from the entrance to overspeeding position. With the speed-measuring snapshot device, it can directly obtain the overspeeding position of vehicle, namely the distance S, the overspeeding velocity of vehicle V_{os} and the speed limit of the section as V_{ls}.

So the time it takes for a vehicle to travel from the freeway entrance to overspeeding snapshot position at the overspeeding speed can be expressed as

$$T_1 = \frac{S}{V_{os}} \tag{1}$$

Assuming the vehicle travel at the speed limit, the time it takes for the vehicle to travel from the freeway entrance to overspeeding snapshot position would be

$$T_2 = \frac{S}{V_{ls}} \tag{2}$$

2.3.2 Determination of Time Interval Speed

Since the interval velocity measurement is to set two adjacent monitoring points on the same road section, the principle is to calculate the average speed of the vehicle on the road section based on the time of the vehicle entering the two monitoring points in succession, and judge whether the vehicle is overspeeding according to the speed limit on this section. With the distance from the later monitoring point passed by the vehicle as the distance between the vehicle to the freeway entrance S, the time that the vehicle passes the later monitoring point is taken as the overspeeding time of the unlicensed vehicle. Likewise, by T = S/V, the time it takes for the vehicle to travel from the freeway entrance to the later monitoring points is calculated.

2.4 Determination of Vehicle Entry Time Period

The acquisition of current system time is implemented by the time() function of the C language, and the calculation and process of all time and data is based on the return value of the function. The current time acquired is the number of seconds starting

from 1 January 1970 to the present. With the aid of tm structure, the date and time are obtained. The definition of tm structure in **time. h** is as follows:

Struct tm{

int tm_sec;	The value range of Second is [0,59].
int tm_min;	The value range of Minute is [0,59].
int tm_hour;	The value range of Hour is [0,23].
int tm_mday;	The value range of the Date within a month is [1,31].
int tm_mon;	Month (starting from January, 0 represents January) has the value range of [0,11].
int tm_year;	The value of Year is equal to the actual year minus 1900.
int tm_wday;	The value range of Week is [0,6], and 0 denotes Sunday.
int tm_yday;	The value range is [0,365], e.g. 0 represents January 1 and so on.
int tm_isdest;	Daylight Saving Time TAG

}

Time point of the vehicle entering into the entry can be confirmed by using time backtracking method based on vehicle-hours of travel and overspeeding time of the vehicle. When the vehicle drives at V_{os} speed, time of the vehicle entering into freeway entrance is M1; when the unlicensed vehicle drives at V_{ls} speed, tiem of the vehicle entering into freeway entrance is M2, thus the time of the vehicle entering into the freeway entrance is M2-M1.

3 Application of Algorithm

3.1 Ordinary Overspeeding

Figure 3 shows that: Time: 2016-03-21 18:24:16, Distance: Beijing-Zhuhai Expressway 809 km, Speed: 107 km/h, Speed limit: 90 km/h, Speed ratio: 18%

Fig. 3 Ordinary overspeeding capture picture

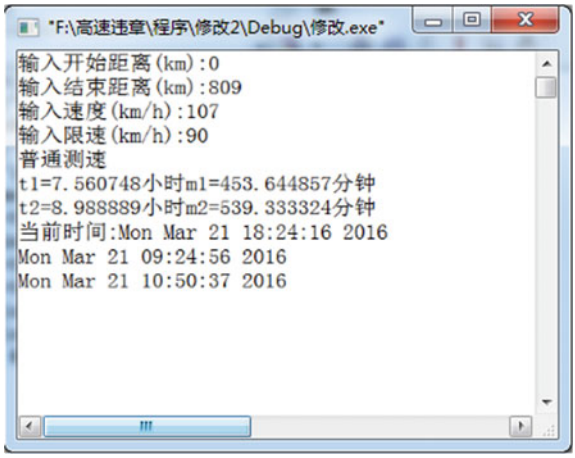


Fig. 4 Operation result

Fig. 5 Ordinary overspeeding capture picture

Figure 4 is a diagram showing the operation result of the overspeed vehicle in Fig. 3.

Figure 5 shows that: Time: 2012-03-08 08:22:39, Distance: Hainan Huandao Highway 37 km, Speed: 156 km/h, Speed limit: 100 km/h, Speed ratio: 56%

Figure 6 is a diagram showing the operation result of the overspeed vehicle in Fig. 5.

3.2 Interval Speeding

Due to the difference between the interval speed measurement and common speed measurement, interval speed measurement means that vehicle speed in certain distance exceeds the speed limit standard of this section, thus the distance that the vehicle drives out of the speed measurement interval is called as the distance from

Fig. 6 Operation result

the vehicle to the freeway entrance. And the time that the vehicle drives out of the section is used as the overspeeding time (Fig. 7).

Time backtracking method is used to confirm the time section for the vehicle to enter into the freeway entrance. Vehicles that pass through the freeway entrance in this time period are sought. Vehicle identification and vehicle landmark identification are used to screen out vehicles at the entrance and ensure vehicle information.

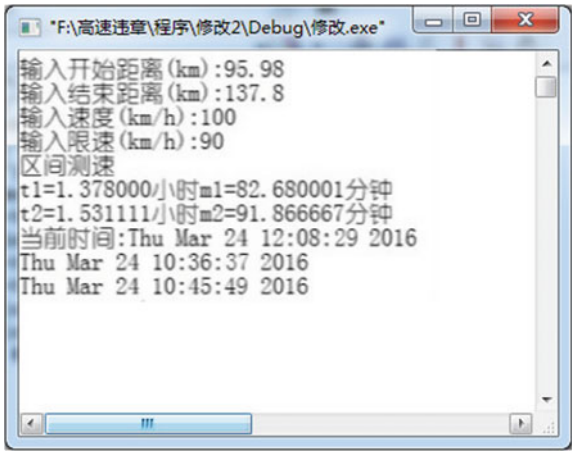


Fig. 7 Interval speed operation result

3.3 Image Screening

When a vehicle passes through the toll station of freeway entrance, an image will be taken. Through the image, it can be clearly observed that which vehicle enters into the freeway from which lane and the vehicle's license plate, model and color.

Time backtracking method can be used to determine the period of time that the vehicle enters the entrance of the freeway is M_2-M_1. Assuming that there are M lanes entering into the freeway from the toll station, the first lane has N_1 vehicles passing through at M_2-M_1 and the second lane at M_2-M_1 has N_2 vehicles, and forth so on. The Mth lane has N_M vehicles at M_2-M_1. Therefore, the total number of vehicles passing through the M_2-M_1 time period is

$$L = \sum_{i=1}^{M} N_i \tag{3}$$

These images are used as image database. Then, image identification is used to do vehicle identification for vehicle images and illegal vehicle images in the database. Model identification can identify brands of illegal vehicles and identify vehicles that have the same model with the illegal overspeeding vehicles. The self-owned learning tool of MATLAB is used for vehicle detection and model identification, as shown in Fig. 8. Vehicle identification based on in-depth study identifies database images. Database images are conducted primary screening to exclude lots of vehicle images and remain vehicles that have the same model with illegal vehicles. The images

Fig. 8 Vehicle Identification Results

screened are used for quadratic search, which searches for vehicle landmarks in the first search, so as to determine vehicle information.

3.4 Verification of Illegal Vehicle Information

The verification principle of illegal vehicle information is similar to the search principle. A speed measurement snapshot device of freeway is used to directly gain the speed, position and image of overspeeding unlicensed vehicles. However, as for verification of illegal information, the distance from the vehicle to the freeway exit is determined by position, so as to gain time of the vehicle reaching the freeway exit, verify the time backtracking through road junction and confirm the overspeeding unlicensed vehicle information based on in-depth study.

4 Conclusion

The overspeeding freeway unlicensed vehicle lookup algorithm based on image processing and image identification screening technology is proposed. The steps and principles of this algorithm are introduced: position and speed of vehicles at expressway are used. Time backtracking method is used to confirm the time section for the vehicle to enter into the freeway entrance. The freeway entrance vehicle image in the time period is used for comparison to confirm illegal vehicle information. The example shows the application of this algorithm. This algorithm applies the time backtracking method of road junction to look for the unlicensed vehicle, but only considers the situation of freeway entrance, instead of vehicle time entering into the service station. There is more than one freeway entrance, so it is necessary to troubleshot the road junction one by one. By aiming at time for vehicles to enter into the service station, a large number of investigations can be used to analyze the relationship between the time when the vehicle enters the service station and the type of vehicle, long-distance or short-distance, etc. The future study will consider service time. Aiming at the problem of needing to investigate information one by one intersection, a more complete vehicle identification method is adopted to improve the accuracy of vehicle identification and increase the efficiency of investigation. A large number of investigations can be used to analyze the relationship between the time when the vehicle enters the service station and the type of vehicle, long-distanceor short-distance, etc.

References

1. Fang A, Miu X (2012) Research on checking system of provincial highway network public security vehicles. Traffic Inf Secur 01:119–124
2. Li Z, Li Y (2012) The development and research status of license plate recognition technology. Technol Inf (5):110–110
3. Ye Q, Cai H, Wang Y (2001) Vehicle type discriminating method based on loop coil and axle load detector. Comput Technol Autom 20(s1):225–227
4. Gu G, Chen Q (2002) An automatic classification electronic charging system based on infrared detection. Traffic Inf Secur 20(6):33–35
5. Lv C, Liu M (2011) Multi-sensor based vehicle classification and identification system for a new type of vehicle. Sci Technol Eng (33):8225–8228,8237
6. Peng B, Zang D (2015) Research on car logo recognition method based on deep learning. Comput Sci 42(4):268–273
7. Deng L, Wang Z (2016) Deep convolution neural networks for vehicle classification. Appl Res Comput 33(3):930–932
8. Song H, Zhang X, Zheng B et al (2018) Vehicle detection based on deep learning in complex scene. Application Research of Computer, 35(04):1270–1273

Study on the Behavior and Psychology of Pedestrian Traffic Violations on the Crosswalk

Ya-xiong Han and Yue-ying Huo

Abstract Pedestrian traffic violations on the crosswalk are common in many Chinese cities. As the weak side within many traffic participants, pedestrian safety is often not guaranteed. In order to reduce pedestrian violations and improve pedestrian safety fundamentally, behavior and psychology of pedestrian crossing the street should be studied and it is addressed in this paper. Firstly, a questionnaire survey was carried out to collect data. Then, frequency and causes of pedestrian violations, pedestrian preferences to crossing facilities were analyzed using mathematical statistics. Based on the analysis above we find that pedestrian violations is related to the job style, and the main reason for pedestrian violations is the psychology of saving, conformity and dependence. Finally, some suggestions about enhancing pedestrian safety education and strengthening pedestrian traffic management are proposed.

Keywords Pedestrian violations · Behavior · Psychology · Design of the crossing facility · Safety education

1 Introduction

Pedestrians are the weak side within many traffic participants, the safety of theirs is often not guaranteed in the face of modern transport. Especially in some intersections and sections of road, lack of effective crossing facilities to segregate vehicle flow and pedestrian flow, pedestrians are often in a dangerous situation. At the same time, the lack of pedestrian moral education has caused the pedestrian violation of the street to be widespread, which has undoubtedly increased the insecurity when pedestrians pass the crosswalk. In the meantime, the traffic control department is

Y. Han · Y. Huo (✉)
Transportation Institute, Inner Mongolia University, No. 24 Zhaojun Road, Yuquan District, Hohhot 010070, China
e-mail: huoyueying2008@163.com

Y. Han
e-mail: hanyx1123@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_53

constantly increasing the punishment for pedestrian violations of the street, which further aggravates the situation of pedestrian violations.

Pedestrian traffic is an important part of slow traffic, more and more urban planning designs are tending to pedestrian traffic gradually. Both the state and society promote the design concept of "people-oriented" actively, the actual demands of pedestrian in the planning and designs of traffic are considered. Thus, in order to reduce pedestrian violations radically and improve the present situation of pedestrian violations, the relevant departments should carry out pedestrian traffic management and safety education effectively.

2 Literature Review

Domestic research on the behavior and psychology of pedestrian violations included the study of pedestrian traffic behavior with different groups, pedestrian traffic behavior model that based on different influence factors, analysis of psychology of pedestrian traffic violations and etc. Wang et al. [1] studied the traffic behavior of middle and primary school students and education of traffic safety, they finally proposed a targeted strategy of traffic safety education for middle and primary school students by investigating the traffic behavior characteristics of middle and primary school students, and then using the mathematical statistics method to study the law of behavior characteristic. Shi et al. [2] studied the knowledge, attitude and behavior(KPA) of traffic safety for middle school students, considered the differences of individuals, they proposed to strengthen the safety education of middle school students. For the model of pedestrian traffic behavior, Guo et al. [3] used logit regression model to analyze the effects of different influence factors on the traffic behavior, and the MNL model of pedestrian behavior is established by using the logit model of pedestrian traffic behavior based on different influence factors [4]. For the study of pedestrian psychology, it is mainly the application of planned behavior theory. Sun [5] discussed the law of psychological awareness and activity of pedestrian violations based on psychological expectation theory. Yuan and Xiao [6] studied the behavior and psychology of pedestrian by establishing the calculation model of recognition process of pedestrian information. In addition, Yang et al. [7] studied the behavior and psychology of pedestrians in Shanghai through the analysis of the questionnaire that based on planned behavior theory.

Foreign research on the behavior and psychology of pedestrian included analysis on characteristics of pedestrian traffic behavior in a specific case, and starting from the psychological characteristics of pedestrian, to improve the traffic environment, then the corresponding solutions was studied. The Israeli author Yagil conducted a questionnaire survey on the influence of the environment and the psychology of pedestrian on the intersections of signal and elaborated the psychology that influenced pedestrian traffic behavior [8]. Hermann and Kamala used discrete choice model to study the choice behavior of drivers and pedestrians meet in the crosswalk, the research has shown that number of pedestrians, the urban scale, vehicle speed, vehicle

queue length and the distance of pedestrians in curbstone will have an effect on pedestrian decision [9]. CJ Hughes, based on psychology, he research on area coloring method of crosswalk and intersection thoroughly in order to improve the safety of pedestrian and provide a good traffic environment for pedestrians [10].

At present, research on pedestrian violations of the street is mostly based on the establishment of the model of pedestrian behavior and use of behavioral observation methods. There are few researches on the psychological study of pedestrian violation behavior, and many are qualitative analysis using some psychological theories or models, which is not accord with the actual psychological status of pedestrian crossing. Therefore, based on pedestrian traffic behavior survey, this paper analyzed the psychology of pedestrian violations from the angle of pedestrian, and put forward some strategies and suggestions to improve the situation of pedestrian violations.

3 The Survey of Pedestrian Traffic Behavior

3.1 Design of Questionnaire

The questionnaire first made statistics on the gender, age and cultural level of pedestrians. For the behavior of pedestrian violations, understanding of violation behavior in our daily, the frequency and cause of violation are included. For the research on pedestrian psychology, pedestrian selection to crossing facilities are studied, which included the pedestrian preference to crossing facilities, the limit of waiting time and the longest distance that acceptable.

3.2 Survey Methods

The survey was based on network investigation, which was supplemented by field investigation. Along with the network popularization, people spend a lot of time on the phone every day, that provides a convenience for data collection, it also reduced the time and cost of the questionnaire. However, the Internet survey has limitations, which is reflected in the small sample size of the elderly and a large sample of young people.

3.3 Sample Characteristics

Of the 207 valid questionnaires, men accounted for 43.96%, and women accounted for 56.04%. In terms of age, the proportion of the people under 20 is 55.07, 39.13% in 20–40 years old, 4.35% in 40–60 years old, and just only 1.45% over 60. Among

them, young people account for the majority of the survey sample, which is related to the frequent use of mobile phones by young people. In terms of cultural level, there are 7.25% of junior high school degree, 56.52% of high school degree, 11.11% of college degree, and 25.12% of undergraduate and above. In terms of daily work style, there are 13.53% of the people chose *Slow motion*, 20.77% of the people chose *Wind and fire* and 65.7% of the people are between the two styles.

4 Analysis of Pedestrian Violation Behavior

In life, pedestrian violations of the street are often related to extent of education, traffic safety awareness and setting of the crossing facility [11]. This section will analyze the reason for pedestrian violations through the understanding of violation behavior, the frequency and cause of pedestrian violations.

4.1 Understanding of Violation Behaviors

The results of the survey are shown in Fig. 1.

In the above statistics, more than 80% people said *Run the red light* and *Turn over and lean in the center of the road* were violations. This is in line with the daily perception that people are cognizant of violation of regulations. In addition, people think that *Not in accordance with the conductor of traffic police* and *Stay in the driveway* were violation behaviors, who tends to have high quality and strong consciousness of traffic safety. Finally, a few of people said that *Not walking on the sidewalk* and *In the absence of road lanes, not by the road* were violation behaviors.

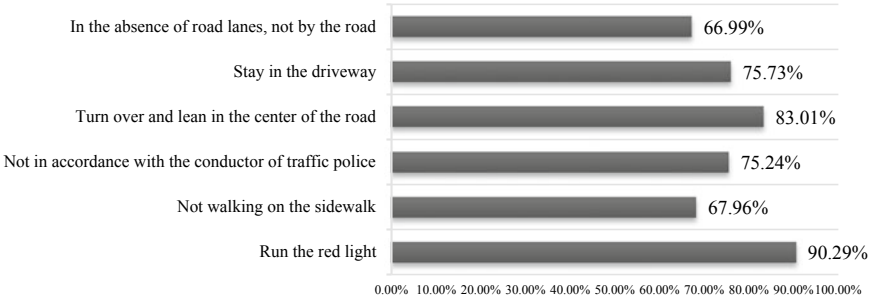


Fig. 1 Pedestrian understanding for common violation behaviors

4.2 Frequency of Pedestrian Violations

The statistics of frequency of pedestrian violations are shown in Fig. 2.

As can be seen from the above figure, pedestrian violations of the street is widespread, the proportion of *Sometimes* and *Regular* is close to 80%, and only 21% people never violate the rules. In the previous section, we discussed the understanding of violation behavior, thinking that the violation was related to such factors as extent of education of pedestrian. For the frequency of pedestrian violations, we will analyze the relationship between frequency and work style of pedestrian.

Through the analysis of the data, the distribution of daily work style for pedestrian is obtained, as showed in Fig. 3.

As can be seen from the above figure, 66% people are between *Slow motion* and *Wind and fire*. This is similar to the *Sometimes* ratio in Fig. 2, whereas most people are between two personalities in reality. These people are in a *tangle* mentality when crossing the street, they are doubts about the violation or waiting for the signal to pass, causing their violation behavior of *Sometimes* [12]. The *Slow motion* people are more conservative when crossing the street, preferring to pass in a safe way, which is close to the proportion of people who never violate the rules in Fig. 2. The *Wind and fire* are more impulsive, considered safety in the street, they would choose to violate occasionally, so the frequency of regular violations is greatly reduced.

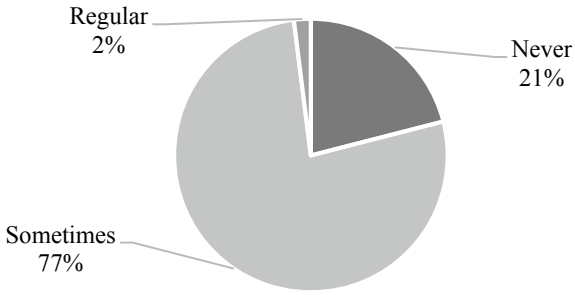


Fig. 2 Frequency of pedestrian violations

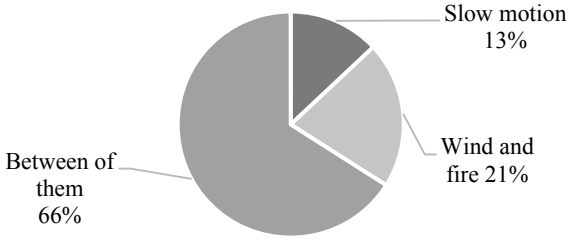


Fig. 3 The daily work style of pedestrian

4.3 The Causes of Pedestrian Violations

For the analysis of the causes for pedestrian violations, psychological factors and traffic environment were studied. The results of the survey are shown in Fig. 4.

Based on the survey statistics of above figure, the reasons for pedestrian violations were analyzed from psychological factors and traffic environment mainly.

From the analysis of psychological factors, there are common psychology of saving, conformity, blind confidence and dependence when pedestrians cross the street. In terms of saving psychology, pedestrians often choose to pass with short walks, close proximity and save energy. The survey found that 61.35% of those who violated the rules chose *In order to catch the time*. Under the action of saving psychology, pedestrians inclined to cross the road, big step to walk. Sometimes they cross the street in dangerous and massive traffic, caused great danger to the security of pedestrians. As for conformity mentality, pedestrians often follow the traffic rules and walk smoothly to through the road when they walked alone considering the weakness of their own strength. But when passers-by gather in the street, someone breaks the street, pedestrians are affected by the people who cross the street, there will be *The law does not punish numerous offenders* and *Follow someone else* mentality. According to the data, 27.54% people had *Follow someone else* conformity behavior. For blind confidence and dependence, the survey found that 30.92% of pedestrians were confident they could pass safely. At the same time, 7.73% of pedestrians had a dependence on the street and they thought that motor vehicles should avoid pedestrians.

From the analysis of traffic environment, urban transport planning in China focus on the motor vehicle, planning and design of pedestrian traffic and its supporting facilities are not enough attention, crossing facilities with low service quality, so pedestrian violations become more serious [13]. According to the data, 25.12% of pedestrians chose *The crossing facilities are not reasonable* and 27.54% of pedestrians chose *It's too much trouble to get to the crossing facilities*.

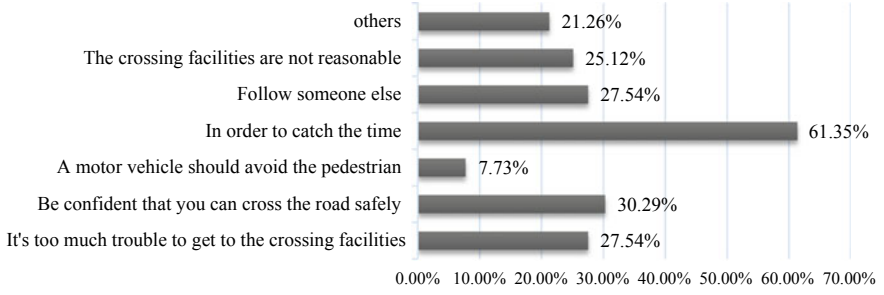


Fig. 4 The main causes of pedestrian violations

5 Analysis of Pedestrian Psychology

5.1 Characteristics of Pedestrian Psychology

Pedestrian psychology has a great influence on behavior, such as *anxious* people and *laid-back* people have a different judgment of passing the street. *Anxious* people are more impulsive, they are more likely to pass at a red light. While *laid-back* person's inner activity is more stable, they tend to pass the street in a safe way, so they seldom violate the rules [14].

This paper analyzes the pedestrian psychology from the perspective of pedestrian selection to crossing facilities, such as the pedestrian preference to crossing facilities, the limit of waiting time and the longest distance that acceptable. It is of great significance to improve the status of pedestrian violations.

5.2 Pedestrian Preference to Crossing Facilities

Pedestrian preference to crossing facilities is shown in Fig. 5. According to the survey data, 43.96% of people choose the flat crossing facility, while 56.04% like the three-dimensional crossing facility. Of the three-dimensional facility, 46.86% are *Footbridge* and 9.18% are *Underground passage* only.

Main considerations for pedestrians chose one crossing facilities to pass the street. There is a survey result shown in Fig. 6.

It is obvious that the main factors that pedestrian consider are security and convenience, which requires that we should pay attention to the practical demand of pedestrians when planning the layout of crossing facilities. In addition to this convenience and safety, there are some people who have higher requirements for the comfort and accessibility of the street. Therefore, it is also important to consider whether the crossing facility can bring us a comfortable environment.

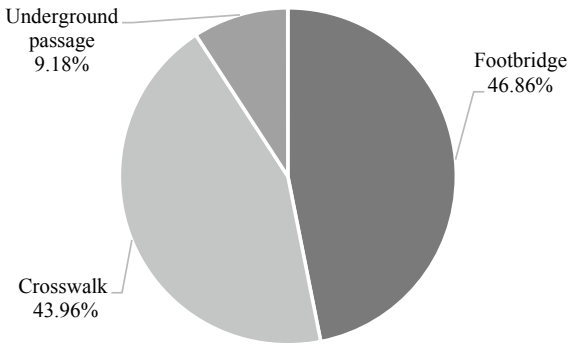


Fig. 5 Pedestrian preference to crossing facilities

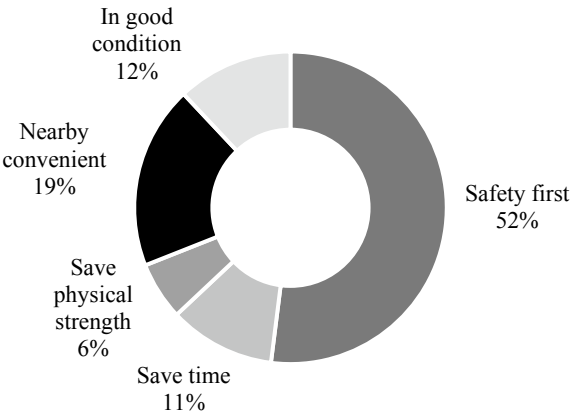


Fig. 6 Main considerations for pedestrians chose crossing facilities

5.3 The Limit of Waiting Time

The limit of waiting time is the maximum acceptable time for a passer-by to wait at the crosswalks, which is discussed in the intersection of signal and no signal control. The specific findings are shown in Figs. 7 and 8.

At the intersection of signal control, pedestrians maintain a restraint on crossing the crosswalk affected by signals. When the waiting time is in the 30 s, pedestrians will violate in the effects of saving psychology. However, when the waiting time exceeds 30 s, the number of pedestrians who cross the crosswalk is reduced, and then increases with the increase of time. When the waiting time exceeds the 80 s, it reaches the peak of pedestrian violations. That is, the limit of waiting time.

At the intersection of no signal control, lack of a stoplight as *constraint* on pedestrians, pedestrians mainly depend on the number of motor vehicles to pass the crosswalk. When the waiting time is in the 30 s, because of no restrictions on signal lights,

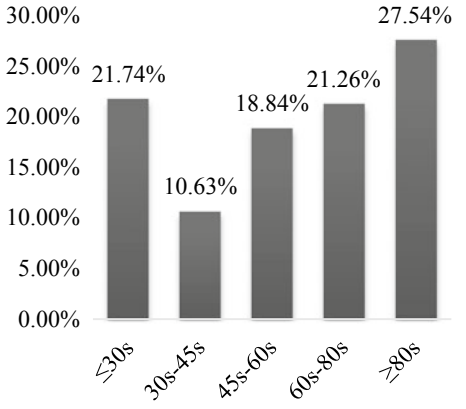


Fig. 7 The limit of waiting time at the intersection of signal control

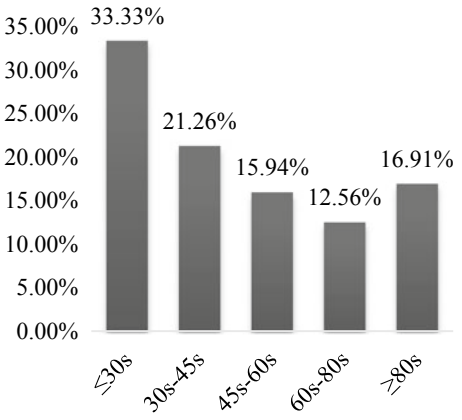


Fig. 8 The limit of waiting time at the intersection of no signal control

many people choose to cross the street, and with the increase of time, the frequency of pedestrian violations is reduced, so the limit of waiting time is 30 s.

When the waiting time of pedestrians is less than the limit of waiting time, pedestrians can pass according to the signal color basically, the control lability of pedestrian traffic flow is better. On the contrary, the control lability of pedestrian traffic flow is poor, and the proportion of pedestrians forced to cross the vehicle flow is very high [15].

5.4 The Longest Distance that Pedestrian Can Be Acceptable

When pedestrians cross the street, they are sensitive to the distance to crossing facility. Too long distance often makes them lose patience and chooses to cross the street [16]. The specific findings are shown in Fig. 9.

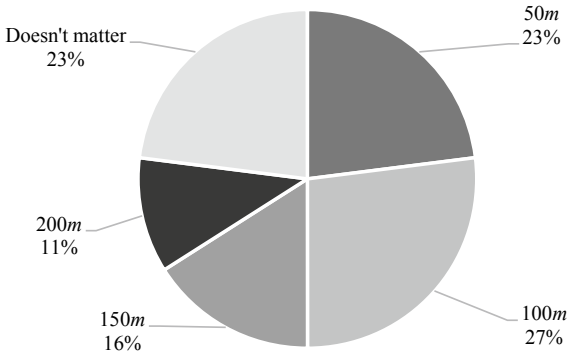


Fig. 9 The longest distance that pedestrians can be acceptable

Table 1 The Longest distance that Pedestrians can be acceptable at different ages

Age range	The distance				
	50 m (%)	100 m (%)	150 m (%)	200 m (%)	Doesn't matter (%)
Below 20	28.07	21.05	14.04	7.89	28.95
20–40	18.52	37.04	13.58	14.81	16.05
40–60	11.11	11.11	44.44	11.11	22.22
More than 60	0	33.33	66.67	0	0

According to the above results, the longest distance that pedestrians can be acceptable is 100 m, which is at 27.05%. When the distance between pedestrian and the crossing facility exceeds this standard, pedestrians often violate the rules. Many factors in pedestrian samples will have an impact on the longest distance, and Table 1 provides a statistical analysis of the distance pedestrians can be acceptable at different ages.

The data in the table shows that the longest distance that pedestrians can be acceptable around different ages are different, which mainly related to behavior characteristics in different age groups, it also related to this article investigation statistics of the sample distribution all ages. But overall, the longest distance that pedestrians can be acceptable focused on 50–150 m, considering the actual situation that people is mainly under the age of 40, the elderly population that over the age of 40 is relatively small, we can determine the longest distance that pedestrian can be acceptable is 100 m.

6 Conclusion

In this paper, in order to reduce pedestrian violations and improve pedestrian safety fundamentally, behavior and psychology of pedestrian crossing the street were studied. Firstly, a questionnaire survey was carried out to collect data. Then, frequency and causes of pedestrian violations, pedestrian preferences to crossing facilities were analyzed using mathematical statistics. Based on the analysis above we find that the main reason for pedestrian violations is the psychology of saving, conformity and dependence. Finally, some suggestions to reduce pedestrian violation are proposed.

- Pay attention to pedestrian traffic safety education

By strengthening the safety publicity and education for primary and middle school students, we can continuously improve the public awareness of pedestrian violations of the street and to improve pedestrian violations. In addition, continue to increase the traffic safety awareness of drivers and pedestrians, strengthen the pedestrian self-suggestion, weakening the pedestrian mentality of savings and conformity through broadcast video of traffic safety regular and out pamphlets safety education [17], to solve the problem of pedestrian violations of the street fundamentally.

- Attention to the planning of crossing facilities

It is of great significance to improve pedestrian traffic behavior in the design of cross-street traffic design to meet the actual demands of pedestrians. Therefore, we must carry on the field investigation to collect the views of pedestrians on the setting of crossing facilities. Improve the utilization rate of crossing facilities and reduce pedestrian violations. The design of crossing facilities could consider the special demands of pedestrians. For example, the height of the steps can be reduced appropriately for the footbridge near primary school. For footbridge and underground passage, increase barrier-free facilities taking into account the universality of the travel group. Thus, the problem that some elderly and the people with disabilities are difficult to pass the street can be solved through.

- Enhance pedestrian traffic management

In addition to necessary fines, people who violates should be forced to participate in service activities. The design of *people-oriented* planning for transport facilities is also essential. For example, the traffic management department could optimize the signal phase matching by investigating the intersection of pedestrian violations. For the special situation of intersections, it is advisable to extend the green time of pedestrians and shorten waiting time for pedestrians to pass the street. Finally reduce pedestrian violation on the crosswalks.

Acknowledgements This work is supported by the National Natural Science Foundation of China (No. 51668048), the Scientific Research Projects in Universities of Inner Mongolia (No. NJZY16022). The authors deeply appreciate the supports.

References

1. Wang X, Wang L, Lin R, Yang D (2010) Analysis of traffic behavior and characteristics of primary school students in Shanghai. J Chin Secur Sci 20(12):122–129
2. Shi Q, Wang X, Yang D, Fang S (2011) Analysis of traffic safety knowledge, attitude and behavior of middle school students. J Chin Secur Sci 21(12):143–152
3. Guo J, Zhang X, Han Y (2012) Computer simulation analysis of pedestrian crossing at signal intersection. Railway Comput Appl 21(11):8–11
4. Zhou Z, Wang W, Ren G, Gong X (2013) The selection model of pedestrian crossing behavior of pedestrian crossing at the intersection of signal control. J SE Univ (Nat Sci Ed) 43(03):664–668
5. Sun S (2007) Research on the behavioral psychology of pedestrian violation at signal intersection. Beijing Jiaotong University
6. Yuan H, Xiao G (2008) Research on pedestrian insecurity based on traffic psychology. J Chin Secur Sci 18(1):20–26
7. Yang Y, Sun J, Li K (2011) Analysis and prediction of the psychological causes of pedestrian violation of the street. J Chin Secur Sci 21(04):20–25
8. Li Zhang (1996) Human error in the complex human system. J Chin Secur Sci 6(6):35–38
9. Koval DO (1998) Human element factors affecting reliability and safety. 406–414
10. Lee JYS, Lam WHIS, Wong SC (2001) Pedestrian simulation model for Hongkong underground stations 554–558

11. Liao L, Zhang X (2012) Analysis of pedestrian traffic characteristics across the street. Urban Constr Theor Res (20):54–57
12. Zhang Q (2014) Research on the behavior of pedestrians in urban roadways. Southwest Jiaotong University
13. Zhang Y (2013) Research on the optimization of pedestrian crossing traffic system at intersections. Chang'an University
14. Duan W, Jiang G (2008) The theory of planning behavior. Prog Psychol Sci 16(2):315–320
15. Wei R, Changqiao S (2006) Preliminary analysis of pedestrian crossing traffic behavior. Road Traffic Safety 6(3):33–36
16. Chen Tuansheng, Fang Y, Yang L (2007) Study on the influencing factors of travel choices of elderly people. J Southwest Jiaotong Univ 5:17–21
17. Pan H, Chen P (2010) Pedestrian violation of the street from the street. Traffic Stand 150–156

Cooperative Optimization of Seat Control and Ticket Price for High-Speed Rail Passenger Transport

Zhen-ying Yan, Fang Gao, Ping-ting Zhang, Yujia Zhang, Hui Liu, Xiao-juan Li and Jian-wei Ren

Abstract With the rapid construction of the high-speed railway in China, the operational efficiency and revenue performance are getting more and more attention from operators. In this paper, the idea of revenue management is adopted to adjust the passenger flow of parallel trains through differential pricing considering the preference of passengers for different parallel trains. Based on utility theory, the Multinomial Logit model is used to describe passengers' choice behavior among parallel trains, and a cooperative optimization model for ticket price and seat control of high-speed rail network with multi-train and multi-stop stations is established. At the same time, the optimization strategy of differential pricing and seat control of parallel trains is obtained. The validity of the model is proved by numerical experiments based on the train of Beijing-Shanghai high-speed railway. The optimization polices of this model in this paper can provide optimization strategy for the revenue management of high-speed parallel trains and enrich the synchronous optimization theory of capacity control and pricing in revenue management.

Keywords Revenue management · Railway transportation · Capacity control · Differential pricing

1 Introduction

In the past decade, the construction of high-speed railways in China has been booming. At the end of 2017, the operating mileage of high-speed railways has exceeded 25,000 km, accounting for more than 66.3% of the total mileage of high-speed railways in the world. But in the operation of high-speed railway, there are still many problems such as bad performance of revenue and unbalanced passenger flow. So,

Z. Yan · F. Gao · P. Zhang · Y. Zhang · H. Liu · X. Li (✉) · J. Ren
Transportation Institute of Inner Mongolia University, Hohhot 010070, China
e-mail: 311987431@imu.edu.cn

Z. Yan · X. Li · J. Ren
Inner Mongolia Engineering Research Center for Urban Transportation Data Science and Applications, Hohhot 010070, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_54

the railway operator in China is actively exploring strategies to improve revenue and operational efficiency. In 2017, some tickets for high-speed railway trains on the southeast coast of China were subjected to differential pricing and discounted strategies for early booking. The development of more sophisticated revenue management technology is of great significance for high-speed rail operation management.

At present, there is less research on revenue management in the field of railway. The existing research focuses mainly on the aviation field, and most studies only focus on dynamic pricing and inventory control. A very small number of studies focus on the collaborative optimization of pricing and inventory control, and it is shown that combining the price with the seat control will significantly improve the operating income [1]. In this paper, we develop the cooperative optimization model for multi-trains with multi-stops to make joint decision of pricing and seat control. Passengers have different preferences for trains at different departure time [2]. We use the Multinomial Logit model to describe the passenger choice behaviors among trains. Based on revenue management theory, the deterministic nonlinear programming with the goal of the total revenue maximization is developed to get policies of differential pricing among trains and seat control.

This article is structured as follows: In Sect. 2, we review the scientific literature of revenue management and position our work. In Sect. 3, we state the problem and develop the optimization model. Then we make the numerical experiments in Sect. 4. In Sect. 5, we conclude the paper.

2 Literature Review

We review the literature from pricing, seat control and joint-decision making. In the field of aviation, Gallego and Van Ryzin [3] based on the strength control theory, proposed a classical GVR model for single product pricing, then many subsequent studies are extended on this basis. Li [4] studied the passenger choice behavior among trains of a single OD, then proposed dynamic pricing model for two parallel trains.

There are many contributions for seat control policy. In the field of aviation, You [5] established a non-linear integer programming model to optimize multi-class seat allocation of a single train with multiple stops. Numerical experiments show that the algorithm is faster than Lingo and Dicopt when there are more than a certain number of segments. But this model doesn't consider passenger choice behavior, and Cooper et al. [6] pointed out that the unsatisfied part of the demand was neglected, which would lead to a reduction in revenue. Later in many network revenue management, customer choice behavior was considered. Chen and Homem [7] adopts discrete customer choice behavior to study network aviation revenue management. In the field of railway, Bao et al. [8] described the theory and method of seat control for railway trains and established a dynamic seat control model. Then Wang et al. [9] considered the quantity constraints of different seat classes and established a stochastic programming model considering passenger choice behavior.

At present, the research on the joint optimization of differential pricing and seat control is mainly focused on the aviation field. Weatherford [10] firstly put forward a revenue management model considering both pricing and seat control. Different levels of price influenced demand and price should be taken as a decision variable. Numerical experiments show that the suboptimal solution without cooperative decision is 3–5 times less than the optimal solution of joint decision. Cizaire [11] studied the joint optimization problem of ticket price and stock control. It is found that the combined optimization increases the revenue compared with the traditional revenue management, and then expands to a multi-stage stochastic model. In the field of railway, Hetrakul and Cirillo [12] integrated customer choice behavior into the revenue management for railway system pricing and seat allocation optimization and a joint decision model is established. But they focused on the optimization of one train with assumption of price changing with booking time.

In summary, domestic and foreign scholars have done a lot of in-depth research on the theory of revenue management, but the theory and method for joint decision-making on pricing and seat control in high-speed trains are not comprehensive enough. The existing research on the joint decision of pricing and seat control mainly focuses on the joint decision of single product or two kinds of products, and there is no research on the joint decision making of high-speed railway with network effect. In this paper, the cooperative optimization strategies for differential pricing and seat control of parallel trains at the network level are studied, and this method can be used as a reference for other industries.

3 Model

3.1 Problem Statement

The operator runs several trains on one High-speed railway line with multi-stops. So every train severs passengers from origin station and destination station, and passengers of the same OD select the train based on their preferences. In the actual operation, the price of different train for the same OD are same. But different trains with different performance, like departure time, running time and so on. Then passengers concentrate a certain train according to utility maximization. So, the operator makes the decision of differential price for trains to adjust the distribution of passenger flow and maximize the total revenue of the High-speed railway line. At the same time, the operator makes the decision of seat allocation for different trains and different ODs. We assume that the demand of OD has been forecasted. And we don't consider refunding and No-show.

3.2 Formulation

3.2.1 Passenger Choice Behavior of Different Trains

In the model presented in this paper, the Multinomial Logit model is used to characterize the passengers' choice of trains. According to literatures, departure time, running time and price are the main factors affecting the passengers' choice behavior. d_{ad}^{j} is the passenger's willingness to pay for the departure time of train j for OD, t_{od}^{j} denotes running time of train j for OD, and $fare_{od}^{j}$ denotes the price of train j for OD. We construct the generalized cost function of the train j from o to d as follow:

$$U_{od}^{j} = \beta \cdot t_{od}^{j} + d_{od}^{j} + fare_{od}^{j} + \varepsilon_{j} \tag{1}$$

where: U_{od}^{j} is the generalized cost of the passenger selected the train j from o to d, β is the time value of the passenger, and ε_j is a random term of unobservable cost. Assuming ε_j for any j obeys independent and identical Gumbel distribution, based on the theory of stochastic utility maximization, the probability of passengers choosing the train j from O to D is as follows:

$$p_{od}^{j} = \frac{\exp \theta U_{od}^{j}}{\sum_{j=1}^{J} \exp \theta U_{od}^{j}} \tag{2}$$

where θ is the positive proportional factor. Note that parameters β, γ and θ need to be estimated from historical data or survey data by the maximum likelihood estimation.

3.2.2 Objective Function

The total demand of passengers from origin o to destination d is defined in this paper as D_{od}, which has been forecasted. α_{od}^{j} is the acceptance rate of booking requirements for train j. So $D_{od} \cdot \alpha_{od}^{j}$ can be explained as the number of passengers who buy the tickets of train j between o and d. The total expected revenue of all the trains running on the line is R, and the ultimate goal is to maximize R. The objective function is expressed as follows:

$$\max \mathrm{R} = \sum_{o=1}^{n-1} \sum_{d=o+1}^{n} D_{od} \cdot \left(\sum_{j=1}^{J} \alpha_{od}^{j} p_{od}^{j} fare_{od}^{j} \right) \tag{3}$$

where α_{od}^{j} and $fare_{od}^{j}$ are decision variables to decide seat allocation and price separately and simultaneously.

3.2.3 Constraints

There are three kinds of constraints to be set in this paper. The first kind is capacity constraints. The seat capacity for section l of train j is represented by C_l^j. It should be guaranteed that the number of seats sold for all products across section l cannot exceed C_l^j:

$$\sum_{o=1}^{n-1} \sum_{d=o+1}^{n} a_{od}^{l} \alpha_{od}^{j} D_{od}^{t} p_{od}^{j} \leq C_l^j, \forall l, j \tag{4}$$

The second kind of constraints is about price. The maximum price can't exceed the passenger's acceptance range and the price stipulated by the management department. The minimum price can't be lower than minimum operating cost. The highest fare is defined as F^+, and the lowest fare is defined as F^-.

$$F^- \leq fare_{od}^{j} \leq F^+, \forall od, j \tag{5}$$

The third kind of constraints is about α_{od}^{j}. The value of α_{od}^{j} is between 0 and 1.

$$0 < \alpha_{od}^{j} < 1, \forall od, j \tag{6}$$

The cooperating optimization model is formulated by (3–5). This model is a deterministic non-linear programming, many existing solvers support this kind model, like lingo.

4 Numerical Experiments

We design numerical experiments based on the high-speed train G143 and G115 running from Beijing to Shanghai. As shown in Fig. 1, four stops are selected: Beijing South, Tianjin South, Nanjing South and Shanghai Hongqiao. t_{od}^{j}, d_{od}^{j}, D_{od} and the current fares are shown in Table 1. We consider that the choice behavior of passengers follows the Multinomial Logit model. According to Zhao et al. [2], We set the parameters d_{od}^{j} shown in Table 1 and $\beta = 32.576$, $\theta = -0.075$. The upper and lower limit of fare is set to 0.5 and 1.5 times of the current fare separately. We

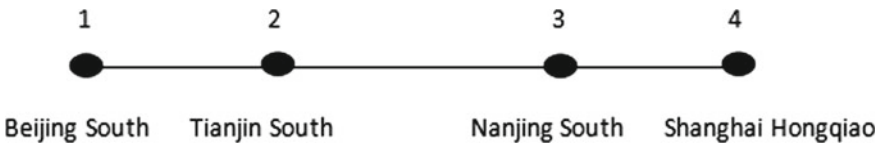


Fig. 1 Train stops

Table 1 Reference data

OD	t_{od}^{j}		Departure time		d_{od}^{j}		D_{od}	$fare_{od}^{j}$
	G143	G115	G143	G115	G143	G115		
1,2	0.6	0.6	7:50	9:20	10.9	16.7	450	54.5
1,3	4.1	4.3	7:50	9:20	86.7	132.8	970	433.5
1,4	5.4	5.8	7:50	9:20	110.6	169.4	1240	553
2,3	3.5	3.7	8:31	9:57	119.2	100.2	530	389
2,4	4.8	5.3	8:31	9:57	152.7	128.7	1078	498.5
3,4	1.3	1.5	11:57	13:42	33.5	21.9	527	109.5

use Lingo 17.0 to solve the model, and the values of α_{od}^{j} and $fare_{od}^{j}$ are shown in Table 2. In order to observe the results of different demand intensities, we test the model with different times of D_{od}.

From the calculation results, we can see that if the demand level is low, all reservation requests are accepted, and the fare is taken to the highest value to obtain the maximum benefit. As the demand level gradually increases, it can be seen from the acceptance rate that the model allocates more seats to the OD with higher earnings performance and optimizes the allocation of train seat resources to different OD. The demand distribution of short-distance ODs is less differentiated, and the demand distribution of some long-haul ODs is more differentiated. The model comprehensively considers the demand distributions of different OD to make differential pricing policies for ODs with significant differences in demand distribution, and same pricing

Table 2 The values of α_{od}^{j}, $fare_{od}^{j}$ and revenue

(*o*,*d*,*j*)	0.1 D_{od}		0.4D_{od}		0.5 D_{od}		D_{od}	
	α_{od}^{j}	$fare_{od}^{j}$	α_{od}^{j}	$fare_{od}^{j}$	α_{od}^{j}	$fare_{od}^{j}$	α_{od}^{j}	$fare_{od}^{j}$
(1,2,1)	1.00	81.8	1.00	81.8	0.00	81.8	0.00	81.8
(1,2,2)	1.00	81.8	1.00	81.8	0.00	50.6	0.00	81.8
(1,3,1)	1.00	665.3	0.00	665.3	0.54	665.3	0.00	665.3
(1,3,2)	1.00	665.3	0.00	539.4	0.00	271.8	0.48	665.3
(1,4,1)	1.00	829.5	0.00	276.5	0.53	829.5	0.48	829.5
(1,4,2)	1.00	829.5	1.00	829.5	0.95	829.5	0.00	829.5
(2,3,1)	1.00	583.5	0.95	583.5	0.00	583.5	0.00	583.5
(2,3,2)	1.00	583.5	0.44	583.5	0.00	583.5	0.00	583.5
(2,4,1)	1.00	747.8	0.00	249.3	0.00	711.8	0.00	249.3
(2,4,2)	1.00	747.8	0.00	747.8	0.00	290.4	0.00	249.3
(3,4,1)	1.00	164.3	1.00	164.3	1.00	164.3	0.00	164.3
(3,4,2)	1.00	164.3	0.45	164.3	0.00	163.3	0.00	164.3
Revenue	582,509.8		914,727.9		978,810.0		978,810.0	

policies for ODs with the smaller difference in distribution. When demand is sufficient, the model allocates seats only for long-distance ODs because of their earnings is best. It is the default that the model doesn't integrate price elasticity of demand in.

5 Conclusions

In this paper, the heterogeneity of train choice behavior for passengers under the multi stop of parallel trains is considered, and the key factor of passengers' preference for departure time is integrated into the model. In order to solve the problem of unbalanced train occupancy, the paper combined the differential pricing strategy and the seat control strategy to establish the optimal cooperative model based on the passenger choice behavior. Taking two parallel trains with four stops as an example, this paper designs numerical experiments with different demand level. The results show that the cooperative optimization model is effective to make joint decisions of pricing and seat control. And the choice behavior of the passengers among the parallel trains has a significant effect on the demand distribution among parallel trains. So the operator can make differential pricing policies for parallel trains to adjust demand distribution and improve revenue performance.

This paper studies the cooperative optimization strategy of differential pricing and seat control under the deterministic demand, but there is often randomness in demand. And the pricing strategy does not take into account the impact of passenger demand changes in the booking period on fares, so the price doesn't change with booking time. Therefor how to integrate dynamic pricing and stochastic demand will be the next direction of research.

Acknowledgements This work is supported by the Research Program of Science and Technology at Universities of Inner Mongolia Autonomous Region (NJZY18012); the National Natural Science Foundation of China (No. 51668048); and the Inner Mongolia Natural Science Foundation (No. 2017BS0501). The authors deeply appreciate the support.

References

1. De BSV (2003) Advances in airline revenue management and pricing. Massachusetts Inst Technol 10–16
2. Zhao P, Li YF, Li B (2017) Study on the train choice behavior for high-speed railway passengers considering the departure time preference. J Beijing JiaoTong Univ 41(6):2–3
3. Gallego G, Van Ryzin G (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. Manage Sci 40(8):999–1020
4. Li YF (2017) Research on passengers' booking behavior under conditions of high speed railway parallel trains. Beijing JiaoTong University, Beijing
5. You PS (2008) An efficient computational approach for railway booking problems. Eur J Oper Res 185(2):811–824

6. Cooper WL, Homem-De-Mello T, Kleywegt AJ (2006) Models of the spiral-down effect in revenue management. Oper Res 54(5):968–987
7. Chen L, Homem T (2010) Mathematical programming models for revenue management under customer choice. Eur J Oper Res 203(2):294–305
8. Bao Y, Liu J, Ma MS (2012) Analysis and Simulation of arrival process of passenger ticket purchase request during ticket Pre-sale period. J Beijing JiaoTong Univ 36(6):27–32
9. Wang X, Wang H, Zhang X (2016) Stochastic seat allocation models for passenger rail transportation under customer choice. Transp Res Part E 96:95–112
10. Weatherford LR (1997) Using prices more realistically as decision variables in perishable-asset revenue management problems. J Comb Optim 1(3):277–304
11. Cizaire C (2011) Optimization models for joint airline pricing and seat inventory control: multiple products, multiple periods. Massachusetts Inst Technol
12. Hetrakul P, Cirillo C (2014) A latent class choice based model system for railway optimal pricing and seat allocation. Transp Res Part E 61:68–83

Influence of Lane Change on Driving Behaviours in Traffic Oscillations Based on Vehicle Trajectory Data from Aerial Videos

Qian Wan, Guoqing Peng, Zhibin Li, Wenyong Li and Qianqian Liu

Abstract Many Car-Following (CF) models and analysis methods have been applied to many practical and theoretical studies, relatively, only a few in Lane-changing (LC) development. This research aims to fill the gap by proposing a new tracking lane-changing trajectory and theoretical method to study date. In this paper, we employed Unmanned Aerial Vehicle (UAV) to record the moving data of the vehicles in Nanjing, China. Based on existing research methods, we study the influence of lane-changing (LC), a comprehensive data analysis indicates that drivers show similarity of their lane-changing habit but with variety, and different drivers' lane-change trajectory data show different lane-change "personality" including aggressive and timid characteristic. By analyzing the data and comparing it with the related research based on NGSIM, we can obtain the corresponding changes in driver characteristics: (i) A timid (aggressive) driver tends to become less timid (aggressive) or convert to slightly aggressive (timid) after experiencing LC; (ii) These changes were systematic and suggest that drivers tend to become more aggressive (characterized by decreasing average time headway after LC) perhaps to prevent another LC occurring. The research conclusions of this paper are similar to those of the existing research results, but also have some innovation points, so it can be proved that the data extraction method and the theoretical analysis method in this study are reasonable and innovative. Therefore, what we found in this paper are significantly helpful to study the characteristics of Chinese drivers, and which have enlightening effect to

Q. Wan · G. Peng (✉) · W. Li
Guilin University of Electronic Technology, Jinjilu #1, Guilin 541004, China
e-mail: 243564846@qq.com

Q. Wan
School of Architecture, Southeast University, Sipailou #2, Nanjing, China

Hualan Design and Consulting Group, 39 Hua Dong Road, Nanning 530011, China

Z. Li
School of Transportation, Southeast University, Nanjing 210018, China

Q. Liu
Guangxi University of Finance and Economics, #100 West Mingxiu Road, Nanning, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_55

intelligent transportation system (ITS), unmanned driving and other new technology application in traffic field.

Keywords Lane-changing analysis · Traffic flow · Trajectory extraction · Unmanned aerial vehicle application

1 Introduction

Classical lane-changing date mainly comes from NGSIM which collected in USA around 10 years ago, however, it exists some limitations to study traffic phenomena in other area to some extent. The influence of lane change on traffic flow is not only individual, especially when there is a large number of traffic flow changing, it will have a considerable impact on the overall operation state.

Sunyu [8] has quantified the interaction between vehicles when changing lanes and proposed a competitive/cooperative model to reproduce the traffic flow in different traffic jams. Wang Xiaoyuan, Xing Li and Wu Fang of Shandong University, from the perspective of the driver's psychological-physical characteristics, use the Analytic Hierarchy Process (AHP) to quantify the successive levels of driver's decision-making thinking. The decision model of driver lane change based on AHP is established. Talebpoura et al. [9] proposed a game theory based change model. This model is applied to the communication between vehicles in 80s.

Wen-Long Jin re simulated the effect of interference lane change by modifying the velocity density relationship, and proposed a dynamic model of lane change traffic based on motion wave. Nagel et al. [6] put forward the famous NaSch model, which describes the acceleration and deceleration of vehicles and the updating process of vehicle position with simple evolution rules, and can reproduce various phenomena in traffic flow. Barlovic et al. [1] Proposed a speed dependent VDA model based on NaSch model, and regarded the speed as a random probability rather than a constant of the slowing process. According to Maerivoet and Moor [3], the differences between different lane changing models are mainly reflected in the requirements of lane changing and the minimum gap between lane changing. Kukida [2] is also based on the NaSch model, and introduces the changing rules considering the speed difference of the near lane to improve it.

There are few literatures on the influence of the change of lane on the traffic flow. At present, the research of lane change is limited to the model of lane change and the influence factors of lane-changing behaviour, and the relationship between lane change and traffic congestion is less considered. Facing the present traffic situation, it is necessary to study the relationship between lane change and traffic congestion.

In this study, we study the entire transition period, consisting of anticipation and relaxation and changes in driver characteristics due to LC. To facilitate this, we develop a new method to systematically process the date collected in Nanjing Kazimen viaduct, and use similar methods as Zheng et al. [10] to analyse the transition process induced by LC and measure its ensuing impact on driver characteristics. The

study results in this paper have certain guiding significance to the actual engineering construction, traffic management, the further research in the driver characteristics and the intelligent transportation system.

This paper is organized as follows. Section 2 describes the data acquisition site, time and method. Section 3 describes the methodologies to process and identify the transition implication of a follower during LC. Section 4 presents the computing and compared results of driver characteristics by applying the similar methods as Zheng et al. [11] model. Finally, Sect. 5 summarizes our findings and discusses future research.

2 Trajectory Data Acquisition

Nanjing Kazimen and Shuangqiaomen viaduct are the southeast of Nanjing interchange hub, from the confluence to the exit ramp about 250 m long. Vehicles intertwine with each other when changing lanes here, as shown in Fig. 1, resulting in a decrease in traffic efficiency and causing congestion. This study selects the section of approximately 300 m from the on-ramp to the off-ramp, and mainly divides the road into 4 sections (the sequence numbers are as follows), the Sects. 1 and 3 are the main lines, and 2 (4) is the section of the on-ramp (off-ramp). The video was recorded by UAV (Unmanned Aerial Vehicle) from 17: 30 to 18:30 (peak hour) on sunny day. In the video recorded, the following charts show the traffic volume of each section.

Table 1 shows that the in main-line (Sect. 1) traffic volume at the entrance is about 3336 veh/h, the on-ramp (Sect. 2) is about 3072 veh/h; the out main-line (Sect. 3)

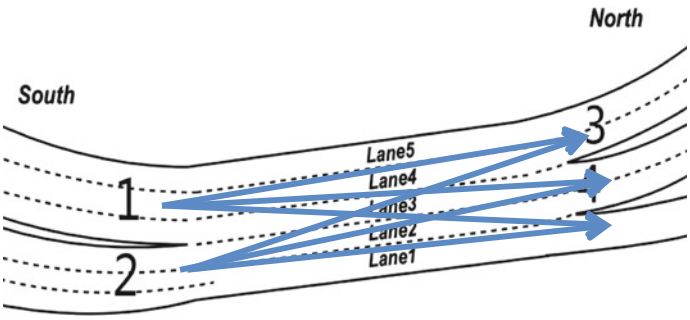


Fig. 1 4 sections are marked on the road in this study

Table 1 4 sections traffic volume

	Section 1 (in main-line)	Section 2 (on-ramp)	Section 3 (out main-line)	Section 4 (off-ramp)
Peak hourly volume (veh/h)	3336	3072	2550	3048

Table 2 Traffic volume between two section

Direction	Volume (veh/h)
Section 1–4	1470
Section 2–3	282

traffic volume at the exit is about 2550 veh/h, and the off-ramp (Sect. 4) is about 3048 veh/h.

As Table 2, by counting the traffic flow in the directions 1 to 4 and 2 to 3, it can be seen that the size of the different direction traffic flow intertwined situation. According to the obtained dates, it can be judged that there are many interweaving vehicles, and combined with the actual situation, it can be qualitatively analysed that the situation of the intersection at the entrance is more serious.

3 Methodology

We extracted 30 groups as samples, and there are 3 vehicles [eader (veh1), lane changer (veh2), follower (veh3)] included in every group. The specific sample requirements are following:

① veh1 and veh3 in the same lane before the veh2 inserting;
② only the veh2 (one car) inserting in this process;
③ veh1–3 in the same lane after the veh2 inserted;
④ Each sample consists of four processes.

(normal driving, anticipation, relaxation, normal driving).

Sample characteristics:

① The sample duration was 15–30 s;
② The samples were sedan cars and jubilee wagons;
③ Speed was lower than 30 km/h (because traffic during the study periods was congested);
④ The vehicles in the sample are located in the main line (between the entrances and exits) and near the ramp.

In this study, 30 vehicles that meet the requirements of analysis were selected to analyse the behaviour of lane-changing. Main line sample vehicles mostly distributed in 4 or 5 lanes (1–3 main lanes are intertwined and vehicles are more discrete, so it is not suitable to be used as a sample). The 2–3 lanes are located near the exit (at this point the fleet is more linear and is suitable for sampling).

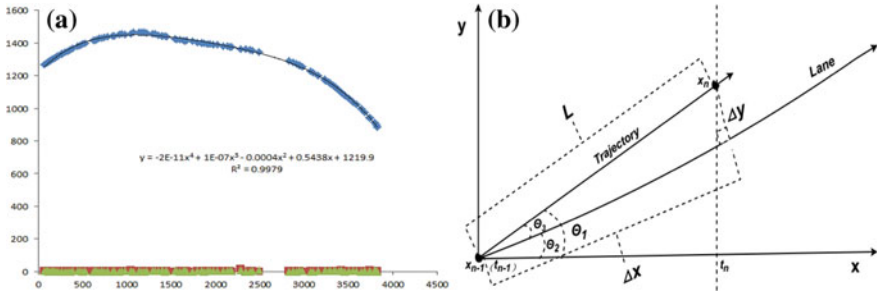


Fig. 2 **a** Actual road in video fitting curve on excel; **b** Coordinate system transformation sketch map

3.1 Road Curve Fitting

Coordinate origin in the upper left corner (MATLAB default setting). The X value increases to the right (Y value increases to the down) in MATLAB picture. The method of road curve fitting is: in MATLAB, taking the points along the road, get the coordinate data of the trajectory along the road, and fit it in Excel. In this paper, the road curve fitting shown as Fig. 2a.

3.2 Coordinate System Transformation and Calculation Process

Because the trajectory of the sample is not straight (the actual lane is curved), it should be projected into a straight line for analysis, and the trajectory that may actually produce a left or right deviation relative to the lane should be projected perpendicularly to the lane direction, as shown in Fig. 2b. Coordinate system is MATLAB coordinate system. X_{n-1} and X_n are the actual position of the vehicle at two times (t_{n-1} and t_n). After correction along the driveway, the actual distance ($\triangle X$) and deviation ($\triangle Y$) from the actual road are obtained (maximum $\triangle Y$ when changing lanes and small fluctuations in the rest of the time). θ_1 is the angle between the actual trajectory direction and the transverse coordinate; θ_2 is the angle between the actual lane direction and the transverse coordinate; θ_3 is the angle between the actual trajectory direction and the lane direction, $\triangle Y$ is close to 0 when vehicle run along the lane direction. The calculation process of the parameters in the diagram is as follows,

$$\theta_3 = \theta_1 - \theta_2 \tag{1}$$

$$L = \sqrt{(X_n - X_{n-1})^2 + (Y_n - Y_{n-1})^2} \tag{2}$$

The frames of each sample video are equal to the number of coordinates obtained. In this case the distance of movement is too small to analysis (coordinate data can be very close or even equal to 0) because of the actual movement of 0.04 s per frame (Tan = 0 may occur when calculating θ). Therefore, this study takes 10 frames as a moving unit for analysis (10 fps = 0.4 s), it means that there are 10 coordinates between X_{n-1} and X_n,

$$tan\theta_1 = \frac{Y_n - Y_{n-1}}{X_n - X_{n-1}} \quad \rightarrow \quad \theta_1 = arctan\left(\frac{Y_n - Y_{n-1}}{X_n - X_{n-1}}\right) \tag{3}$$

Now it need to calculate the θ_2. Because the speed of the sample is less than 30 km/h (8.3 m/s), so $(X_n - X_{n-1}) < 3.32$ m (8.3 * 0.4). In the actual road, it can be thought that the moving distance (3.32 m) is a approximate straight line. That is, the X_{n-1} to X_n sections trend to be straight line. Therefore, the midpoint coordinates $X_{(2n-1)/2}$ between X_{n-1} and X_n are taken, and by substituting it into the derivative calculation $\left(Y'(x)\right)$ of fitting curve function, θ_2 can be obtained by solving inverse tangent,

$$tan\theta_2 = Y'\left(X_{(2n-1)/2}\right) \quad \rightarrow \quad \theta_2 = arctan\left(Y'\left(X_{(2n-1)/2}\right)\right) \tag{4}$$

$$\triangle X = L * cos(\theta_1 - \theta_2) \quad \triangle Y = L * sin(\theta_1 - \theta_2) \tag{5}$$

When △X and △Y are obtained, the graph can be generated in Excel (in Sect. 4 it would be shown).

3.3 *Observation of Lane-Changing Induced Transition*

Based on Newell's CF theory and the research of Zuduo Zheng et al. [10], as shown in Fig. 3, this paper determines the starting point of anticipation (P_1), the starting point (P_2) of relaxation and the end point (P_3) of the relaxation. In this study, t and d are used to represent the time headway and space headway, and it can reflect a transformation relationship between the two vehicles in time and space. In the above chart, t_1 and d_1 (t_2 and d_2, t_3 and d_3) are calculated from the trajectory dates of veh1 and veh3 (veh1 and veh2, veh2 and veh3) before (after) inserting. The parameters t and d can maximize the correlation between the two vehicle trajectories. It is also noted that the trajectory of vehicle 3 during the insertion of veh2 represents a transitional period and should be excluded from parameter estimation. According to MA and Ahnan [5], the impact of LC is usually less than 30 s, and the trajectory dates within 15 s are excluded. Therefore, in this paper, the time span of sample is during 15–30 s.

In the figure, region 1 (region 4) shows the following behaviour in stable traffic flow before (after) insertion. Region 2 and region 3 are the areas where the traffic behaviour changes due to the insertion of the veh2. Region 2 is anticipation and

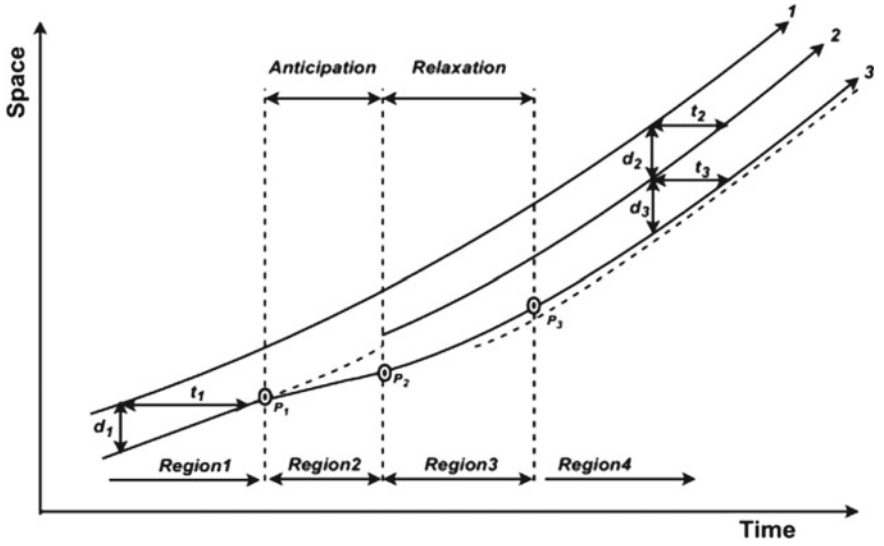


Fig. 3 Observation of lane-changing induced transition to illustrate the anticipation and relaxation, vehicle 1 (leader), vehicle 2 (lane changer), and vehicle 3 (follower)

deceleration process for veh3 before veh2 inserting, to provide more space for veh2 inserting. Region 3 represent the relaxation process between veh2 and veh3. The results of Zuduo Zheng's study show that the two processes are different, but the differences are quite small (There are similar results in this study), so the two processes can be regarded as a unified transition process.

In the above figure, anticipation (relaxation) is the interval between P_1 and P_2 (P_2 and P_3). If vehicle 2 does not insert, the theoretical trajectory of vehicle 3 can be obtained by shifting the trajectory of vehicle 1 to t_1 and d_1 (the dotted line before the insertion of vehicle 2 in the figure). And the trajectory of the actual vehicle 3 starts to deviate from this theoretical trajectory at P_1, because vehicle 3 starts predicting that vehicle 2 may soon be inserting, so decelerating. Therefore, this study defines the P_1 as the star point of anticipation. Similarly, the theoretical trajectory of vehicle 3 after the veh2 inserting is the dotted line behind P_2 in the diagram. Vehicle 3 is not on the theoretical trajectory, indicating that the driver's characteristics have changed (Zuduo Zheng). P_2 is the insertion time point for vehicle 2, that is, the time point at which the trajectory of vehicle 2 was first recorded in the target lane (which can be seen in the trajectory diagram of the following samples). P_3 serves as the end point of relaxation, that is, the starting point for the rebalancing of three vehicles.

4 Results

In this section, we adopt the above analysis method to obtain date (t and d), and used similar techniques as the Zheng et al. [10] to analyse and compare the results with their research.

4.1 Statistics

Some rules can be found from 30 samples taken in this experiment, in which P_1 is judged by the time point when the trajectory of 3 vehicles in ΣXi-T diagram begins to deviate from the theoretical trajectory. P_2 combines two kinds of judgment methods: One is that the actual trajectories diagram of 3 vehicles (X-Y diagram), the trajectory of veh2 is about to coincide with the trajectory of veh1 and veh3 (insertion position shown in the following figure); the other method is the end point of Y fluctuation of veh2 in Y-T diagram. The point P_3 is judged by the time point at which the space distance of three vehicles in ΣXi-T diagram tends to be stationary.

Table 3 is a list of the time headway and space headway of the 30 samples in this study following.

4.2 Regressive Effect of LC on Driver Characteristics

4.2.1 Observation

As described in Sect. 3, (t_1, d_1) and (t_3, d_3) should be similar if the driver behaviour experiencing LC remains unchanged. Our analysis imply that this is often not the case (seen in Table 3, the values of t_1 and d_3 are significantly different).

In order to further investigate the changes in driver characteristics before and after insertion, the changes ($d_3 - d_1$ and $t_3 - t_1$) are plotted against the amount of deviation from "average" driving behaviour before LC ($d_1 - \bar{d}$ and $t_1 - \bar{t}$), where $\bar{d}$ and $\bar{t}$ are average values of d and t prior to LC. As shown in Fig. 4, these plots reveal seemingly linear trends with negative slopes. We show the plots of the 30 examples. The trends imply that most drivers with t larger (smaller) than $\bar{t}$ would adjust his/her driving by reducing (increasing) t after LC. The same interpretation applies to d. We characterize a driver as "timid" if $t_1 > \bar{t}$ and/or $d_1 > \bar{d}$. A driver is characterized as "aggressive" if $t_1 < \bar{t}$ and/or $d_1 < \bar{d}$. Note that the trajectory of a timid (aggressive) driver would lie below (above) the trajectory of an average driver. Using $\bar{t}$ and $\bar{d}$ provides a straightforward and robust way to distinguish timid/aggressive drivers: (i) it is derived from Newell's CF model and thus maintains the theoretical consistency with other analyses in our study; and (ii) it is not sensitive to the individual measurement errors and consistent across sites.

Table 3 d illustrates the space headway; t illustrates the time headway

Sample number	veh1 and 3 (before insertion)		veh1 and 2 (after insertion)		veh2 and 3 (after insertion)		ΔP_1P_2 (Anticipation) (s)	ΔP_2P_3 (Relaxation) (s)
	d_1 (m)	t_1 (s)	d_2 (m)	t_2 (s)	d_3 (m)	t_3 (s)		
1	13.1	4.2	6.5	2.5	4.2	1.8	4	10.4
2	10.4	2.5	21.2	2.3	11.1	1.4	5.5	3.4
3	10.1	2.6	12.8	1.9	12.5	2.5	5	4.6
4	13.7	2.1	13.0	1.9	13.6	2.1	4.3	3.4
5	14.1	2.7	15.2	1.6	20.0	2.2	5	5
6	12.3	2.7	19.2	1.9	11.3	1.5	4.6	5.8
7	11.1	2.3	17.3	2.2	11.0	1.6	5.4	3.7
8	11.6	3.3	18.3	1.9	12.0	1.7	4.9	5.7
9	12.7	2.4	7.6	1.6	10.5	2.1	1.6	6
10	10.5	3.1	8.3	1.3	15.0	2.5	7.7	7.5
11	15.5	2.5	8.0	3.9	6.4	2.4	3.4	4.2
12	9.6	2	5.2	1.7	8.9	2.8	6	2.8
13	24.7	3.6	8.8	2	11.4	2.5	2.9	5.9
14	13.1	2	10.1	1.6	11.3	2	5.6	2.8
15	9.2	3.4	14.8	5.1	4.1	2	8.2	8.6
16	10.6	3.5	6.5	4.4	7.5	4	7.4	7.4
17	10.2	2.6	4.6	1.4	9.1	2.8	4	4.4
18	8.7	2	7.8	1.3	7.5	1.5	8.7	2.5
19	11.3	2	8.1	1.5	11.5	2.4	5.8	5
20	15.1	3.3	9.9	1.6	16.8	1.7	3.6	6.4
21	11.2	1.6	10.3	1.7	12.8	2	2	5.2
22	9.9	2.4	7.6	1.6	14.9	3.6	3.1	6.1
23	11.7	3.2	21.3	4.9	2.9	1.1	6.8	6.4
24	9.9	2.6	16.5	4.1	5.2	1.4	2.9	7.5
25	13.6	2.8	6.2	1.1	12.4	2.1	5	5.4
26	7.9	2.8	14.8	2.4	6.0	1.2	2.8	7.2
27	16.7	2.6	12.2	1.9	12.3	2	3.7	5.9
28	11.1	2.7	6.4	1.8	10.0	3.1	4	5.6
29	10.1	2.4	3.2	0.6	19.3	2.5	3.2	8.6
30	15.7	2.9	6.8	1.7	9.6	2.8	3.4	8.2
Average	12.2 (3.16)	2.7 (0.56)	10.9 (5.11)	2.2 (1.11)	10.7 (4.10)	2.2 (0.66)	4.7 (1.76)	5.7 (1.88)

The value is respective standard deviation in parentheses

4.2.2 Evaluation on Model

Base above analysis, we preliminarily postulate that a LC has a regressive effect on driver CF behaviour: a timid driver with larger response time (t) and/or minimum spacing (d) tends to become less timid and an aggressive driver less aggressive after a LC. The extent of a behavioural change depends on the amount of deviation from the "average" driving behaviour before experiencing a LC. In testing the regressive

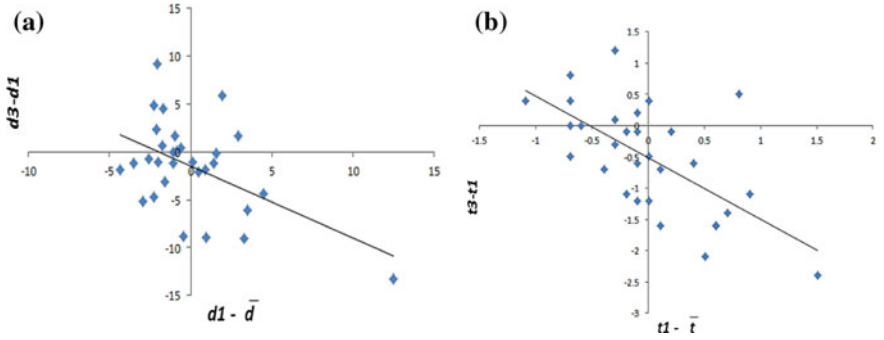


Fig. 4 Change in driver characteristics versus deviation from average driving behaviour; **a** $d_3 - d_1$ versus $d_1 - \bar{d}$ for followers; **b** $t_3 - t_1$ versus $t_1 - \bar{t}$ for followers

effect of LC, we consider the following linear framework as in Zheng et al. to keep the model parsimonious:

$$t_3 - t_1 = \alpha(t_1 - \bar{t}) + \beta \tag{6}$$

$$d_3 - d_1 = \alpha(d_1 - \bar{d}) + \beta \tag{7}$$

where α and β represent independent variable.

We actualized the Univariate Linear Regression Analysis. Both Tables 4 and 5 summarize the estimated parameters and supporting statistics for followers. In our study, the R^2 (in Table 4) ranging from 0.26 to 0.41, and the Zheng et al. ranges 0.30–0.51. Notably, all linear models are statistically significant at the 99% confidence level ($p < 0.01$). Considering that individual driving behaviour is modelled, these R^2 values are quite reasonable.

However, the β exists significantly different between Zheng et al. ($\beta \approx 0$ in t) and our research ($\beta = -0.517$ in t).The difference triggers an important influence on following analysis. To gain more insight on this relationship, as in Zheng et al., Eq (6) is rearranged to obtain

Table 4 Summary of modelling results: followers using 30 samples in this study

Site	Dependent variable	Variable	Coefficient	*p*-Value	Overall goodness of fit	
					p-Value	R^2
Nanjing Kazimen viaduct (China)	$t_3 - t_1$	α	−0.980	<.01	<.01	0.41
		β	−0.517	<.01		
	$d_3 - d_1$	α	−0.756	<.01	<.01	0.26
		β	−1.480	0.06		

Table 5 Summary of modelling results: followers using data from NGSIM by Zheng et al. [11]

Site	Dependent variable	Variable	Coefficient	*p*-Value	Overall Goodness of Fit	
					p-Value	R^2
I-80 (US)	t_3–t_1	α	−0.670	<0.01	<0.01	0.41
		β	−0.0188	0.88		
	d_3–d_1	α	−0.819	<0.01	<0.01	0.30
		β	−1.281	0.24		
US-101	t_3–t_1	α	−0.776	<0.01	<0.01	0.36
		β	−0.0461	0.73		
	d_3–d_1	α	−0.782	<0.01	<0.01	0.51
		β	−0.853	0.42		

$$t_3 - \bar{t} = (\alpha + 1)\left(t_1 - \bar{t}\right) + \beta \tag{8}$$

Based on the modelling results shown in Table 4, $-1 < \alpha < 0$ (or $0 < \alpha + 1 < 1$), the value of $\left(t_1 - \bar{t}\right)$ in our test mostly range −0.6 to 0.6 (shown in Fig. 8b), and $\beta = -0.517 < 0$. Therefore, the $\left(t_3 - \bar{t}\right)$ theoretical calculated results could be either positive or negative numbers (however, these results are closed to 0). Obviously, this relationship implies that a timid driver, who is characterized by $t_1 - \bar{t} > 0$, may become less timid $\left(t_1 > t_3 > \bar{t}\right)$ or slightly aggressive $\left(t_3 - \bar{t} < 0\right)$ after experiencing LC. Similarly, an aggressive driver $\left(t_1 - \bar{t} < 0\right)$ may remain aggressive $\left(t_3 - \bar{t} < 0\right)$ after experiencing LC or convert to slightly timid $\left(t_3 - \bar{t} > 0\right)$. The same results are obtained using d because the estimated coefficients based on d are similar to those based on t.

However, what we found and our above conclusions differ from the study of Zheng et al. [11]. They concluded that the impact of LC on the immediate follower's driving behaviour is not strong enough to convert a timid driver to an aggressive one or vice versa. From our perspective, the reason why the different conclusions appeared maybe due to the different data collected in different areas (Our study date collected in Nanjing, China; Zheng et al. used the date from NGSIM which collected in USA). Therefore, the differences from driving characteristics in two countries have resulted in the diversities in our study, but the diversities are not significant. On the other hand, we both have a same conclusion that LC "neutralizes" the immediate follower's behaviour by encouraging a timid (aggressive) driver to become less timid (aggressive) because $0 < \alpha + 1 < 1$.

4.2.3 Macroscopic Performance

The macroscopic implication of the time headway changes in driver characteristics due to LC can be seen in Fig. 5, we exhibit a few examples of t distribution in each

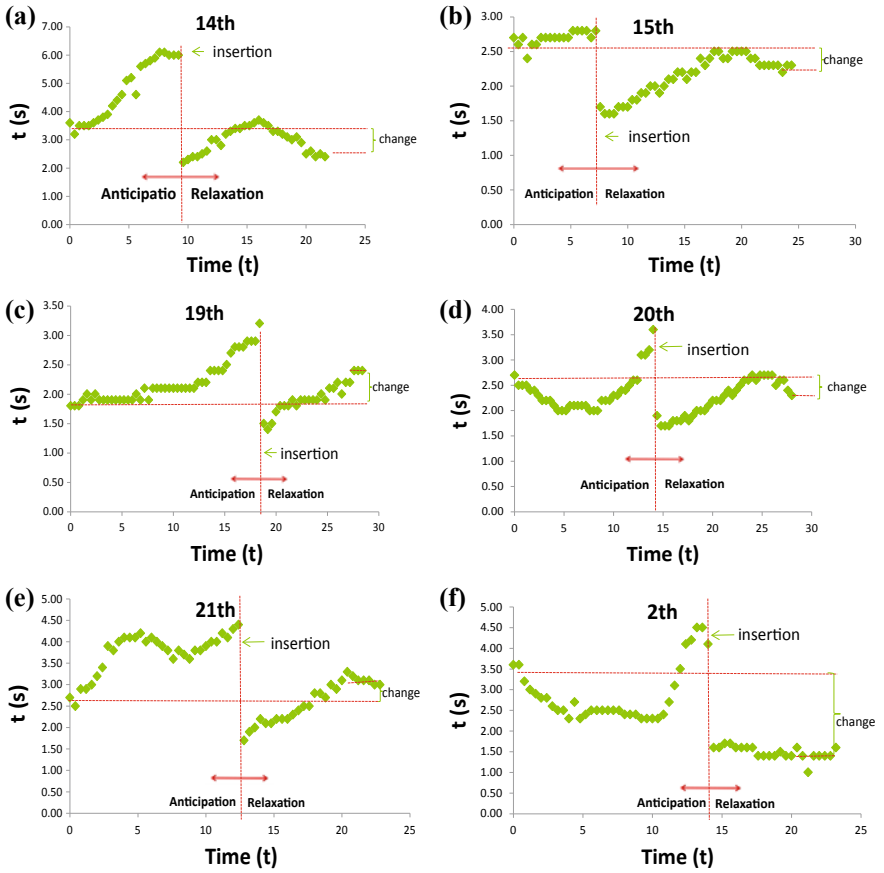


Fig. 5 Temporal evolution of average t in each group. **a–f** respectively represent the group number is 14th, 15th, 19th, 21th, 22th and 2th

group to show that the amplitude of fluctuation of t change and the transition process in driver characteristic.

In this study, we find the phenomenon that the samples whose t reduction ($t_3 - t_1 < 0$) are more than the samples whose t increase ($t_3 - t_1 > 0$) in 30 samples, from which we could conjecture that most driver tend to closer the vehicle in front. We provide following analysis to further validate. The Fig. 6 exhibits the evolution of the t across all the followers in 30 samples in this study. Note that the average t values are calculated based on veh1 and 3 before the anticipation process, and based on vehicles 2 and 3 after the relaxation process. The average t in this paper before anticipation is stable at about 2.69 s (see in Table 6); it changes over time during the anticipation and relaxation processes; then it eventually converges to another average

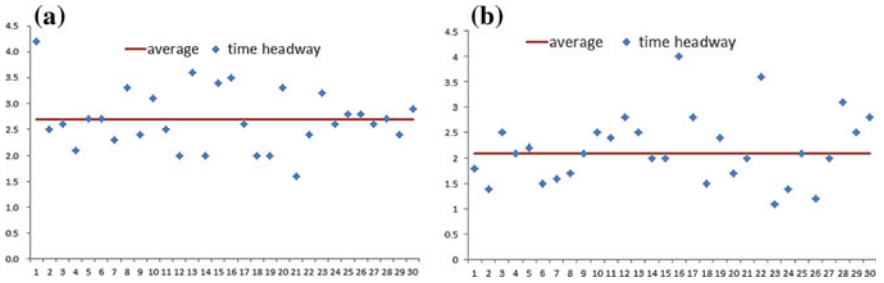


Fig. 6 t distribution of 30 samples in this study. **a** t_1, before anticipation **b** t_3, after relaxation

Table 6 Compared with the research of Zheng et al. [11] (date from NGSIM), the results in equilibrium t shown in the following table

Date (area)	Average t_1 (s) (before anticipation)	Average t_3 (s) (after relaxation)	Rate of decline (%) $\left(\frac{t_1-t_3}{t_1}\right)$
NGSIM (I-80, USA)	1.50	1.30	13%
NGSIM (US-101, USA)	1.50	1.41	5.7%
Our experiment (Nanjing Kazimen, China)	2.69	2.18	18.9

value about 2.18 s after relaxation. The 18.9% decrease in equilibrium t is statistically significant at a 0.05 level and implies a systematic change in driver behaviour after LC.

Obviously, the Table 6 indicate that both t_1 and t_3 larger than the results of Zheng et al. [11], at the same time, our decrease rate (18.9%) is also larger than I-80 (13%) and US-101 (5.7%). Thus, it is notable that the effect of LC is more significant compared with NGSIM in USA, and it imply that Chinese driver would shorten more t (time headway) to close to the vehicle in front. We speculate that because of the more complex road conditions in China, the drivers intend to reduce the chance of being inserted again, so this change in drive character is more obvious.

Both Zheng et al. [11] and our study indicate that in general, drivers become more aggressive (smaller t) after experiencing LC. This performance is consistent with the study result in Sects. 4.2.1 and 4.2.2, and suggests that Newell's assumption of constant t and d should be relaxed to capture time-dependent driver characteristics, especially around major disturbances such as LC.

5 Conclusion

5.1 Summaries of Results

In this paper, we used unmanned aerial vehicle (UAV) to collect video date of the Katzimen viaduct in Nanjing, China. We used road curve fitting, coordinate system transformation and date calculation in MATLAB to obtain the t (time headway) and d (space headway), shown in Table 3.

Based on Newell's CF theory [7] and Zheng et al. [11] research, we analyse t and d in this paper. Our findings imply that a driver corrects himself/herself to drive more "normally" after experiencing LC,

① The change in t or d is linearly related to the deviations from the average driving behaviour (represented by average $\bar{t}$ and $\bar{d}$);
② A timid driver $(t_1 - \bar{t} > 0)$, may become less timid $(t_1 > t_3 > \bar{t})$ or convert to slightly aggressive $(t_3 - \bar{t} < 0)$ after experiencing LC. The similar conclusion applies to aggressive driver $(t_1 - \bar{t} < 0)$;
③ These changes were systematic and suggest that drivers tend to become more aggressive (characterized by t decrease) perhaps to prevent another LC occurring, in other word, most followers intend to be closer to the car in front. The changes in drive character are more obvious than the researches which date from NGSIM.

Therefore, the above findings shed light on transient traffic and variable driver characteristics induced by LC, which will help better describe important traffic phenomena. For example, LC is one of the major factors for capacity drop. Thus, the LC-induced impacts quantified in this study will be useful for estimating/predicting the amount of capacity drop with better accuracy.

Especially in China, the LC researches based on vehicle trajectories from video recorded by UAV are considerably few. Many investigators and scholars study LC or CF (car-following) by using NGSIM which collected in America around 10 years ago. As we all know, the characteristics of drivers and road conditions in China are significantly different to USA. Thus, we need to establish corresponding database and use the traffic data according to actual conditions to study the traffic phenomenon. In this paper, we provide a new method and technique to analyse and reveal more phenomena in traffic flow.

5.2 Deficiency and Future Research

(i) The number of samples in this study is relatively small. Due to the limited duration of video recording and the small number of vehicles suitable for the analysis, only 30 samples were selected for the analysis. In future studies, the length of the video should be increased in order to extract more samples.

(ii) The track algorithm in this paper is more dependent on the stable external environment. In the night, rain and snow weather, and when the video definition is not high, the trajectory extraction will be greatly interfered, and the efficiency will be high or low. Therefore, the stability of the algorithm needs to be further studied, and the background extraction algorithm needs to be further improved to promote the efficiency and robustness.

(iii) If we need to establish the database, we should increase the sample size and increase the vehicle trajectory data under different traffic conditions, such as different periods of time, different sections of road, different types of vehicles and different lanes.

Acknowledgements The authors appreciate the funding support from the Innovation Project of Guangxi Graduate Education (YCSW2018146), the National Natural Science Foundation of China (51508122, 51478113, 51508094), Guangxi science and technology projects (1524800210, Guike-AB16380280, Guike-AB17292087), the Natural Science Foundation of Guangxi (Grant No. 2015GXNSFBA139216), the Natural Science Foundation of Jiangsu (BK20150612), Shanghai Rising-Star Program (16QB1403000), Shanghai Urban-Rural Development Transportation Talents Special Funds, as well as The Scientific research project of Chinese National Ministry of Housing and Urban-Rural Construction (2017-K2-009). Scientific Research Project of Nanning University (2018XJ39).

References

1. Barlovic R, Santen L, Schadschneider A (1998) Metastable states in cellular automata for traffic flow. Eur Phys J B 5(3):793–800
2. Jin WL (2010) A kinematic wave theory of lane-changing traffic flow. Transp Res Part B 44:1001–1021
3. Kukida S, Tanimoto J, Hagishima A (2011) Analysis of the influence of lane changing on traffic-flow dynamics based on the cellular automaton model. Int J Mod Phys C 22(3):271–281
4. Maerivoet S, Moor BD (2005) Cellular automata models of road traffic. Physics Reports 419(1):1–64
5. Ma T, Ahn S (2008) Comparisons of speed-spacing relations under general car following versus lane-changing. Transp Res Rec: J Transp Res Board 2088:138–147
6. Nagel K, Schreckenber M (1992) A cellular automaton model for freeway traffic. J Phys I 12(2):2221–2229
7. Newell GF (2002) A simplified car following theory: a lower order model. J Transp Res Part B 36(3):195–205
8. Sun D, Kondyli A (2010) Modeling vehicle interactions during lane-changing behavior on arterial streets. Computer-Aided Civil and Infrastructure Engineering 25(8):557–571
9. Talebpoura A, Mahmassania HS, Hamdarb SH (2015) Modeling lane-changing behavior in a connected environment: a game theory approach. Transp Res Part C: Emerg Technol 59:216–232
10. Zheng Z, Ahn S, Chen D, Laval JA (2011) Freeway traffic oscillations: microscopic analysis of formations and propagations using wavelet transform. Transp Res Part B 45(9):1378–1388
11. Zheng Z, Ahn S, Chen D, Laval J (2013) The effects of lane-changing on the immediate follower: anticipation, relaxation, and change in driver characteristics. Transp Res Part C 26:367–379

Discriminant Model of Driving Distraction During Mobile Phone Conversation Based on Eye Movements

Lian Xie, Min Duan and Wenyong Li

Abstract In order to investigate the characteristics of drivers' eye movements during distracted driving caused by mobile phone conversation and establish a driving distraction discriminant model, a driving simulation experiment was conducted. The eye movement index data were collected by eye tracker under different traffic scenes which include normal driving and perform simple or complex conversation secondary task on the urban road and freeway, then the variance analysis was used to analyze the characteristics. Finally, according to the characteristics of drivers' eye movements, a driving distraction discriminant model based on fisher discriminant analysis was constructed for different road types. The ANOVA results showed that the effectiveness of road type and conversation task on the cumulative proportion of the driver's focus on the area of interest in the front road is not statistically significant. However, the average duration of the driver's attention under urban road scene is significantly higher than that of the freeway, and with the increasing of difficulty of driving task, the average duration of attention increased significantly. In addition, the road type and conversation task significantly influenced the change range of pupil area. The accuracy rate of the discriminant model is 75.2% for the driving distraction on urban roads, and 78.3% for the distraction on freeway.

Keywords Traffic safety · Distraction · Fisher's linear discriminant · Eye movement · Driving simulation

L. Xie · W. Li (✉)
School Architecture and Transportation Engineering, Guilin University of Electronic Technology, Guilin 541004, China
e-mail: traffic@guet.edu.cn

L. Xie
Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan 430063, China

M. Duan
School of Mathematics & Computing Science, Guilin University of Electronic Technology, Guilin 541004, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_56

1 Introduction

Some previous studies have shown that drivers need to continuously process information from external environment during driving, execute secondary task will occupy some brain resources, resulting in reduced attention resources allocated to the driving task, which will increase driver's driving risk [1]. American 100-car naturalistic driving study shows that 78% of traffic accidents are related to the driver's lack of attention to the road ahead, and nearly 93% of rear-end collisions are related to distracted driving [2]. Similar conclusion was drawn from a traffic accidents investigation in Germany that 65% of the drivers did not recognize the front information or even the information was not obtained timely which caused the accident [3]. Therefore, judging the distraction state of driving is of great significance for reducing traffic accidents caused by distraction and improving the level of road traffic safety.

According to a network survey, in 2014, 60% of drivers in China had used mobile phones during driving. The use of mobile phones during driving not only causes a certain degree of negative impact on driving performance, such as a decrease in lane keeping ability and speed maintaining ability, but also reduces the driver's ability to perceive external stimuli [4]. Yekhshatyan's study found that cognitive distractions caused by mathematical calculations or hands-free conversations can make drivers react to dangerous reactions slowly [5]. In fact, not only visual distraction will affect the driver's visual behavior, cognitive distraction and visual distraction will have a certain negative impact on the driver's visual behavior [6]. However, there are few studies on eye movement characteristics of drivers under cognitive distracted driving conditions.

In order to study discriminant model of distracted driving, a simulated driving experiment was conducted and mobile phone conversation sub-tasks was used to induce cognitive distractions. The driver's eye movement parameters under different test scenarios were collected through eye tracker equipment, and the influence of different degrees of cognitive distraction on the eye movement behavior of drivers was analyzed. Based on the eye movement characteristics, a discriminating model of driving distraction was established to provide a reference for quantification of driving distraction.

2 Methodology

2.1 Apparatus

The equipment included a driving simulator and an eye tracker. The three-screen driving simulator with three degrees of freedom consists of steering wheels, clutches, brakes, throttles, manual transmission joysticks and other components. It can provide various types of scenes such as freeway, mountain road, and urban road, with a good sense of realism and immersion. The Arrington Portable Binocular Movement

Fig. 1 Driving simulator and eye tracker

Instrument was used to collect the eye movement parameters such as fixation time, scan, blink, pupil and so on with the sampling frequency of 120 Hz (Fig. 1).

2.2 Subjects

A total of 16 students and teachers from Guilin University of Electronic Technology were recruited as subjects, including 13 males and 3 females. The subjects were aged from 22 to 27 years old, with an average age of 23 years and an average driving age of 1.37 years. Subjects were required to be healthy, their visual acuity or corrected visual acuity was required to be 0.8 or more, and their driving experience was more than 1 year.

2.3 Driving Secondary Task

Previous studies have shown that the use of mathematical calculation problems of different difficulty can effectively change the cognitive distraction of drivers [7]. Therefore, we take digital mental arithmetic as a secondary task of driving, including 1-digit arithmetic calculation and 2-digit arithmetic calculation which generated two different levels of difficulty for mobile phone calls, simple and complex, respectively.

2.4 Procedure

The 2 (road type) × 3 (driving task) repeated measures design presented two types of road: urban road, freeway, and three types of driving: normal driving, make a simple call when driving, make a complex call when driving. The road type and driving task were the within subjects factors. The dependent variables were ye movement parameters such as gaze duration, cumulative gaze time ratio, pupil area rangeability, and error rate of mental arithmetic.

The experimental procedure was as follows:

(1) Before the test, subjects were asked to complete a questionnaire providing their basic information: age, gender, driving age, driving mileage, etc.
(2) The staff explained the test procedures and ensured that all subjects understood the test requirements and contents before practice. All subjects were required to complete 30 min adaptive simulation driving.
(3) The car seat was required to be adjusted to maintain a comfortable state. Participants were fitted with an eye tracker and the eye movement was calibrated according to the 9-point method. Then record the pupil area of the driver in a non-driving state.
(4) According to road type and driving task type, there are six scenarios in this experiment, which takes about 15 min for each test, and the rest time was 5 min after completing one driving scene. Subjects need to answer 100 questions of one-digit mental arithmetic for the simple mobile phone call driving task through hands-free phone, while the complicated driving sub-task asked driver to answer 100 two-digit mental arithmetic. The staff recorded the correct rate of answers. Six trial scenarios were conducted in random order.
(5) At the end of the experiment, collect laboratory equipment and backup experimental data.

3 Results

Since this experiment mainly studies the driver's attention to the road area in front, the eye tracker supporting software View Point was used to divide the road ahead area into areas of interest, which is called the front area. Selected eye movement indicators for analysis include the cumulative gaze time ratio, fixation duration, and pupil area rangeability. Among them, the cumulative gaze time refers to the total of the driver's gaze time in a certain area of interest [8], which is used to measure the total time spent by the driver in processing a certain area of interest to reflect the driver's attention to the area. The duration of fixation refers to the duration of the center of sight of the visual axis when the driver is looking at the object [9]. The long fixation time indicates that the observer is difficult to extract and process the information, but there is no definite relationship with the value of the gazing object

itself. The pupil area mainly represents the driver's degree of tension and is also an important indicator reflecting the state of attention of visual information. When the psychological workload is relatively high, the magnitude of pupillary changes is also greater [10].

3.1 Analysis of Proportion of Cumulative Gaze Time

Calculate the ratio of the driver's accumulated gaze time to the front area under normal driving and different difficulty calls. During driving, the average gaze time ratio of the driver to the front road area is 85.57%. There is little change in the cumulative gaze time ratio between different road types and different hands-free calls, as shown in Fig. 2.

Take 0.05 as a hypothesis test to analyze the level of significance, and perform analysis of variance (ANOVA). The results of ANOVA showed that there was no significant effect of road type on the cumulative gaze time ratio in the front area ($F = 1.586$, $P > 0.05$). There was no significant difference in the proportion of cumulative fixation time between normal driving and different difficulty conversation tasks ($F = 1.933$, $P > 0.05$). The interaction between road and communication factors was not significant ($F = 1.136$, $P > 0.05$) (Table 1).

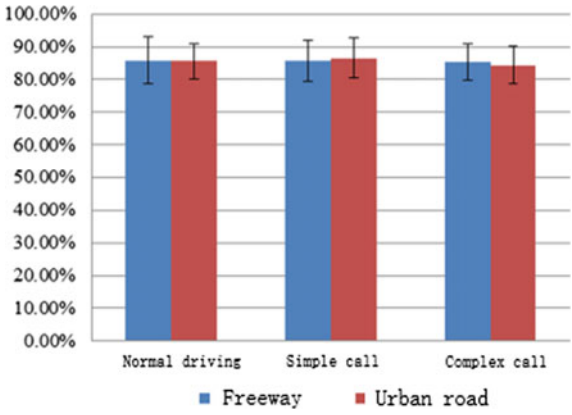


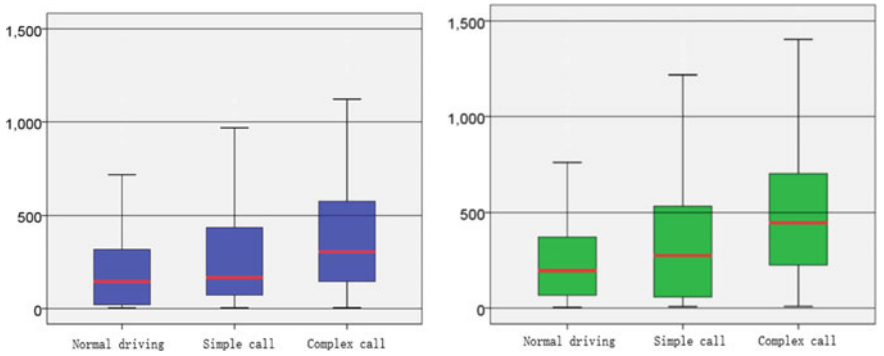
Fig. 2 Histogram of cumulative gaze time in front area

Table 1 Results of cumulative gaze time variance analysis

Independent variables	df	F	P
Road	1	1.586	0.224
Call	2	1.933	0.181
Road * Call	2	1.136	0.349

Table 2 Analysis of variance results of gaze duration

Independent variables	df	F	P
Road	1	79.081	0.000
Call	2	82.115	0.000
Road * Call	2	0.936	0.469



(a) Freeway (b) Urban Road

Fig. 3 The driver gaze duration box chart

3.2 Analysis of Fixation Duration

The data of driver's gaze duration under different test scenarios were extracted and the descriptive statistics and analysis of variance were performed. The results of ANOVA showed that the road type has a significant effect on the gaze duration of the driver ($F = 79.081$, $P < 0.05$). There was significant difference in gaze duration between normal driving and two different secondary task ($F = 82.115$, $P < 0.05$). The interaction between road type and call type was not significant ($F = 0.936$, $P > 0.05$) (Table 2).

The traffic environment under the urban road scene is relatively complex. The driver's averages gaze duration on the gaze point was 365 ms, which was significantly higher than the average gaze duration of the highway which was 279 ms. With the increase of the difficulty of driving task, the average gaze duration increases. The driver reduces scan behavior and the probability of sight shifting between regions, as shown in Fig. 3.

3.3 Analysis of Pupil Area Rangeability

The pupil area of the driver was recorded and compared with that of the driver in the non-driving state, the change value of pupil area was obtained. The results of

Table 3 Results of variance analysis of pupil area variation

Independent variables	df	F	P
Road	1	4.075	0.029
Call	2	15.484	0.000
Road * Call	2	0.991	0.42

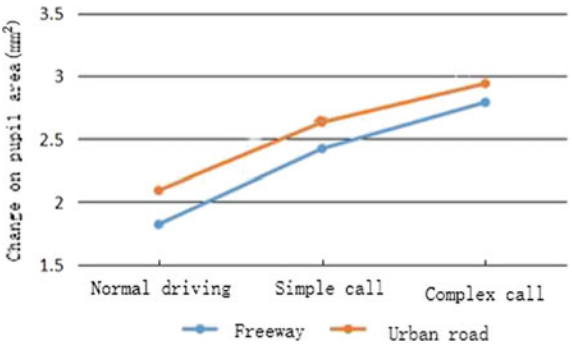


Fig. 4 Fold line diagram of pupil area rangeability

variance analysis for the value of pupil area rangeability showed that the type of road has a significant effect on the change of pupil area ($F = 4.075$, $P < 0.05$). There was significant difference in pupil area between normal driving and different difficulty task ($F = 15.484$, $P < 0.05$). The interaction between road type and driving task is not evident ($F = 0.991$, $P > 0.05$) (Table 3).

Figure 4 showed that the mean change of pupil area in urban road environment was greater than that in freeway (2.55 mm^2 vs. 2.34 mm^2). The change of pupil area increased as driving difficulty increased from non-call task to complex call. Pupil area represents the degree of tension of the driver. When the psychological workload of the driver was relatively high, the change range of pupil surface increased accordingly. The change of pupil area indicated the difficulty of driving on urban roads is relatively high, and the communication task leads to the increase of psychological tension of drivers.

4 Driving Distraction Discriminant Model

In this study, the secondary task of answering mathematical calculation questions by hand-free call was used to induce distraction driving. The average error rate of drivers under simple and complex calls was 4.51 and 16.05%. Therefore, according to the mental arithmetic error rate, the driver's distraction state can divided into three levels: no distraction (normal driving), general distraction, and severe distraction. The results of eye movement analysis showed that there were significant differences in gaze duration and pupil area between normal drivers, simple calls and complex calls.

It means the secondary task could induce distracted driving. Therefore, the Fisher discriminant analysis was used to establish the linear combination of the driver's eye movement parameters to judge the driver's distraction state in the urban road and the freeway environment.

The basic principle of Fisher discriminant analysis is to project all kinds of independent variables in high-dimensional space into low-dimensional space, so that the coincidence of each class in low-dimensional space is minimized [11]. This method is relatively simple and suitable for discriminating the small number of samples. Fisher discriminant analysis assumes that the discriminant function is a linear function. According to the existing sample information, the discriminant function is trained to determine the coefficient value of the independent variable in the discriminant function, and finally the discriminant function relation is determined. Suppose that the discriminant function is:

$$D = u_0 + u_1 X_1 + u_2 X_2 \cdots + u_n X_n \tag{1}$$

In the formula, D is the indicator of driver distraction evaluation; u_0 is a constant term of discriminant function; u_1, u_2, …, u_n is the coefficient of the discriminant function; X_1, X_2, …, X_n for explanatory variables; Gaze duration in this case (X_1), Pupil area rangeability (X_2), Cumulative gaze time ratio (X_3).

In this paper, six test scenarios are designed for two different types of road: urban road and expressway, and three different types of secondary tasks: normal driving and simple call, complex call. Ten samples were extracted from each test scenario according to time interval 1 min, and a total of 16 * 2 * 3 * 10 = 960 samples were collected. Then randomly selected 70% of the sample as the training set and 30% of the sample as the test set, and 3-folder cross validation was used to test algorithm accuracy. The classification function coefficients of different driving distraction levels in urban road environment are obtained as follows:

$$\text{Severe distraction: } D_1 = -21.880 + 4.231X_1 + 4.721X_2 \cdots + 28.307X_n \tag{2}$$

$$\text{General distraction: } D_2 = -15.646 + 2.040X_1 + 3.979X_2 \cdots + 24.670X_n \tag{3}$$

$$\text{Normal driving: } D_3 = -10.512 + 1.189X_1 + 3.414X_2 \cdots + 19.291X_n \tag{4}$$

The classification function coefficients of each driving distraction level in freeway environment are as follows:

$$\text{Severe distraction: } D_1 = -22.523 + 16.233X_1 + 2.445X_2 \cdots + 28.859X_n \tag{5}$$

Table 4 Fisher discriminant analysis accuracy list

Label	Prediction result			
	Severe distraction (%)	General distraction (%)	Normal driving (%)	Summation (%)
Urban Road				
Severe distraction	**80.6**	13.1	6.3	100.0
General distraction	27.5	**55.0**	17.5	100.0
Normal driving	0.0	10.0	**90.0**	100.0
Overall accuracy	75.2			
Freeway				
Severe distraction	**85.0**	15.0	0.0	100.0
General distraction	9.4	**63.8**	26.9	100.0
Normal driving	0.0	13.8	**86.3**	100.0
Overall accuracy	78.3			

Bold indicates rate of correct recognition

$$\text{General distraction: } D_2 = -14.641 + 4.780X_1 + 2.297X_2 \cdots + 26.445X_n \quad (6)$$

$$\text{Normal driving: } D_3 = -11.762 + 2.952X_1 + 1.439X_2 \cdots + 25.515X_n \quad (7)$$

The discriminant accuracy of the degree of distraction in different road types was shown as shown in Table 4. According to Table 4, overall, the discriminant model has an accuracy of 75.2% for driver distraction level discrimination in urban road environments. Among them, the discriminating rate of the two states without distraction and severe distraction was high, exceeding 80%. As for the freeway environment, the accuracy of the judgment of the driver distraction level is slightly higher, reaching 78.3%. And the discriminating rate of the two states without distraction and severe distraction was high, exceeding 85%.

5 Conclusion

Based on the simulated driving platform, the eye movement data of the driver on different road types and driving tasks were collected. Through the analyses of eye movement index, following conclusions were obtained. There is no difference in the

proportion of cumulative gaze time to the forward road, when drivers experienced different road types and different call tasks. However, the average gaze duration in urban road was significantly higher than that in freeway. With the increase of the difficulty driving task, the mean gaze increased significantly and the driver's scan behavior decreased accordingly. Road type and conversation tasks have a significant effect on the change of pupil area. The result of pupil area rangeability indicated that the difficulty of driving on urban roads is relatively high and the conversation secondary tasks can lead to psychological stress.

Based on the characteristics of gaze duration, change of pupil area and cumulative gaze time ratio under distracted driving, a distraction driving discriminant model was established by Fisher discriminant method. The accuracy of urban road distraction discrimination is 75.2% and that of freeway driving distraction is 78.3%.

In this study, the characteristics of eye movement under distraction driving are explored through simulated driving experiments, and a driving distraction discriminant model is established based on the characteristics of driver eye movement. However, the influence of age, experience, sex and other individual differences on the discriminant model was not taken into account. Because of the difference between simulated driving and natural driving, the experimental results may deviate from the actual situation. Therefore, it is necessary to carry out natural driving experiments to improve the research accuracy and reliability.

Funding The authors acknowledge the support from the National Nature Science Foundation of China (nos. 61963011, 5177051327, 51678460, U1664262, 71861006) and the Innovation and Entrepreneurship Training Program for Undergraduates (201810595188).

References

1. Jin LS, Li KY, Xian HC, Gao LL (2014) A study of secondary driving tasks on safety. J Transp Inf Saf 32(5):7–12, 19
2. Dingus TA, Klauer SG, Neale VL, Petersen A, Lee SE, Sudweeks J et al (2006) The 100-car naturalistic driving study: phase II—results of the 100-car field experiment. U.S. Department of Transportation, National Highway Traffic Safety Administration. National Highway Traffic Safety Administration, Washington, DC
3. Graab B, Donner E, Chiellino U, Hoppe M (2008) Analyse von Verkehrsunfällen hinsichtlich unterschiedlicher Fahrerpopulationen und daraus ableitbarer Ergebnisse für die Entwicklung adaptiver Fahrerassistenzsysteme. Tagung aktive Sicherheit, München
4. Ma Y, Fu R (2015) Research and development of drivers visual behavior and driving safety. China J Highw Transp 28(6):82–92
5. Yekhshatyan L (2010) Detecting distraction and degraded driver performance with visual behavior metrics. The University of Iowa, Iowa City
6. Liu ZF, Fu R, Cheng WD, Wu FW (2015) Overview of researches on drivers' visual distraction and cognitive distraction. Chin Saf Sci J 25(7):29–34
7. Massel L, Harbluk J (2006) The impact of performing cognitive tasks on drivers' braking behavior. In: Proceeding of the Canadian multidisciplinary road safety conference XII, pp 1–8
8. Li PF, Wang DH, Liu DB, Wang JJ (2010) Effect of using cell phone while driving on mental workload and driving behavior. J Transp Inf Saf 28(4):103–107

9. Jiang QY, Li J, Cheng PP (2016) Investigation of the safety awareness of the influential factors among the construction workers. J Saf Env 16(6):174–178
10. Arien C, Jongen EMM, Brijs K et al (2013) A simulator study on the impact of traffic calming measures in urban areas on driving behavior and workload. Accid Anal Prev 61(6):43–53
11. Xu CC, Liu P, Wang W, Jiang X (2012) Discriminant analysis based method to develop real-time crash indicator for evaluating freeway safety. J Southeast Univ (Nat Sci Ed) 42(3):555–559

Study of a New Method of Traffic Organization in Reconstruction and Extension of Chang-Zhang Expressway

Yazhen Chen

Abstract Through the study of a new method of traffic organization to reconstruction and extension plan for the reconstruction and extension of the Chang zhang Expressway, from the aspects of program design, emergency response and security measures in the implementation process, the system summarizes the advantages and disadvantages of the transportation organization management for the reconstruction and expansion of the Chang zhang Expressway, and provides reference cases for the subsequent reconstruction and extension of similar highways.

Keywords Chang zhang expressway · Reconstruction and extension · Four-lane traffic flow · Traffic organization

Nanchang to Zhangshu expressway (hereinafter referred to as the "Chang zhang expressway") route length 86.545 km. Chang zhang expressway reconstruction project in jiangxi province is the first large complex highway reconstruction project. in 2012, the traffic flow are more than 45,000 pcu/d, overloading and the surrounding road network traffic low density, There are basically no parallel or connected roads. due to the reconstruction during the period of construction of traffic organization scheme, project management, maintenance, events, condition of emergency treatment etc. are lack of experience, so the reconstruction and the construction of traffic organization and management has brought great difficulty and challenge. The project is With the "three guarantees" as the core concept of highway reconstruction and expansion of traffic organization management [1], from Chang zhang expressway traffic organization plan formulation, this paper analysis the problems in the practical engineering system is summarized, and provide reference for subsequent similar highway reconstruction projects.

Y. Chen (✉)
Wuhan Zhongjiao Traffic Engineering Co., Ltd, Wuhan, Hubei, China
e-mail: 78590823@qq.com

CCCC Second Expressway Consultant Co. Ltd, Wuhan, Hubei, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_57

1 The Present Situation and the Reconstruction and Extension Plan of Chang Zhang Expressway

The whole line of Chang zhang expressway adopts the way of "widening and reconstruction at two sides", and the whole process is changed and expanded according to the eight-lane highway standard. All need to be removed to build 25 new span bridge, new or modified eight Interchanges, joining together the main bridge is about 63, the project itself the complex technology, construction period to maintain traffic flow to traffic is difficult. Comb through the system, the project done to heavy and difficult engineering considerations, such as: the medicine lake used on both sides of the separation of the new scheme of the large bridge, reduces the difficulty of traffic organization.

The regional road network is mainly composed of G105 and G320, and the G105 is parallel to this project. Most of the roads are covered with mud and gravel, and the carrying capacity is limited. The G320 is far away from Chang zhang expressway, The service level is first-class highway, in 2012, the average mixed traffic volume of road section is 12,200 pcu/d, which can assume partial transit traffic flow, but can not afford the transit traffic flow of the road.

2 Division of the Traffic Organization and the Staged Target

The traffic organization design of changzhang expressway reconstruction and extension has adopted the traffic organization plan of "two-way four-lane pass-through" with "road network induced shunting and the parallel construction of the right and left sections of the road section". The programme is mainly introduced as follows:

Stage 1 (10 months): Finish the roadbed to the top of the road bed; the 25 span bridge completed to be removed to rebuilded; Construction of the upper and lower part of the new bridge on both sides of the main line bridge; Connectivity expansion completed (Table 1).

The main objective of the transport organization is:

① The roadbed width hardly affects the operation of the main line vehicle, the car speed limit is 80 km/h, and the speed limit of the cart is 60 km/h without diversion.
The main goal of traffic organization in this stage is to ensure the orderly construction, and avoid the influence of normal traffic flow.

② The new construction and demolition of the bridge to ensure the safety of the principle, as far as possible to the two-way 4 lane normal passage. At this stage some sections are closed for short periods of time.

Stage 2, 3 (about 14 months): old road reconstruction of new and old road splicing; Some sections of the new Bridges are constructed with temporary pavement,

Table 1 The road conditions of different stage of the traffic organizations

Stage	Road condition
Stage 1	The construction of the construction of the new bridge and the lower structure of the main line is completed. (Restricted 4 lane traffic)
Stage 2	New road construct to the surface of the flexible base; finish the new bridge construction. (Restricted 4 lane traffic)
Stage 3	Complete the old road reconstruction; Main line bridge old bridge reinforcement reconstruction. (Restricted 4 lane traffic)
Stage 4	Complete the new and old road splicing and paving, the main line of the new and old bridge splicing and bridge paving and transportation facilities construction. (Locally restricted four-lane, partial eight-lane traffic)

main line bridge old bridge reinforcement and reconstruction, new old bridge splicing. Construction of pavement and bridge surface layer; Construction of the central divider.

The main objective of traffic organization in this stage is:

① Traffic transfer is maintained in two-way 4 lane, as part of the road base is spliced to the flexible base construction to occupy the old road emergency lane, In the condition of ensuring that the vehicle has at least 0.5 m lateral width, the inner lane and the middle divider are required to be divided into 0.25 m respectively for lateral width and lateral width respectively.
The main goal of traffic organization in this stage is to ensure safe and orderly construction under the four-lane traffic flow.

② After construction of half road surface layer, reset the sign and marking for the next phase of traffic, between the two lanes to set up temporary isolation anti-collision facilities such as water horse, at this stage the traffic organization's main goal is to ensure road safety and improve the traffic capacity.

Stage 4 (6 months): all the traffic facilities are under construction, the traffic is open, and the two-way traffic is eight lanes.

3 Traffic Organization Scheme

3.1 Roadbed Road Works

Stage 1 External isolation grid; On both sides of the extension (one-side width 7.5 m), the roadbed is filled to the top of the bed; the traffic is maintained on the old road (Fig. 1).

Stage 2 Each section is divided into four parts according to the demarcation and layout of the main structures along the line, and each part is constructed at the same time, and the construction operation section is shown as below.

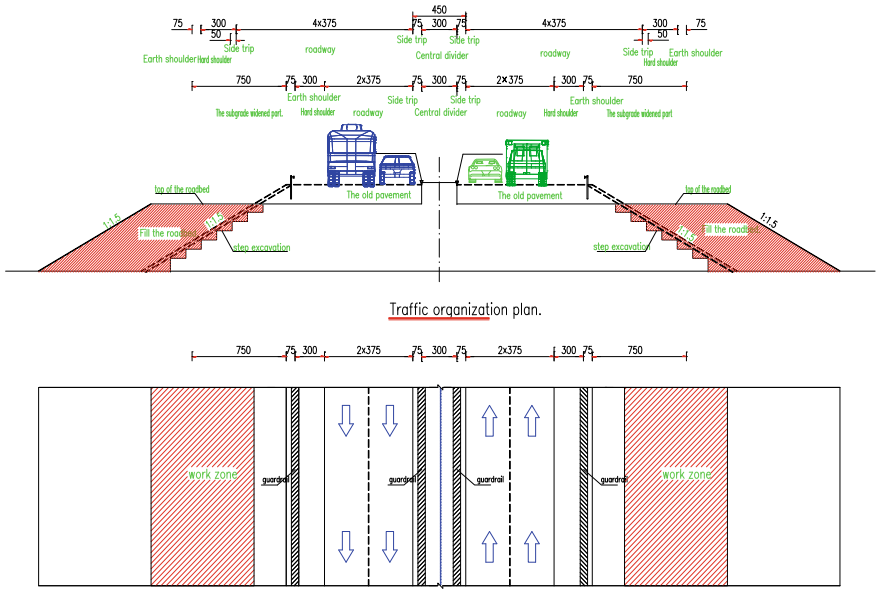


Fig. 1 Stage 1 subgrade construction traffic organization (unit: cm)

The road works are completed in parallel with the cycle of each section (Fig. 2).

Step 1 The two sides of area 3 are constructed to the flexible base, embankment splicing built on both sides of the road to the flexible base (with the old road surface at the same elevation level), set up temporary isolation facilities between new and old pavement, this phase should be the inside of the car lane is set to 3.5 m guarantee in the absence of the hard shoulder of the second half of the picture two lanes outside each have the lateral width of 0.5 m, the traffic at old roads maintains four lanes (Fig. 3).

Step 2 Repeat the first step, The two sides of area 4 are constructed to the flexible base, the hard shoulder one invisible 3 m range, embankment splicing built on both sides of the road to the flexible base (with the old road surface at the same elevation level).

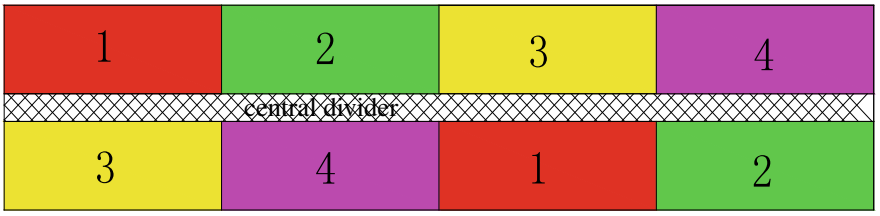


Fig. 2 Section division of section in road construction

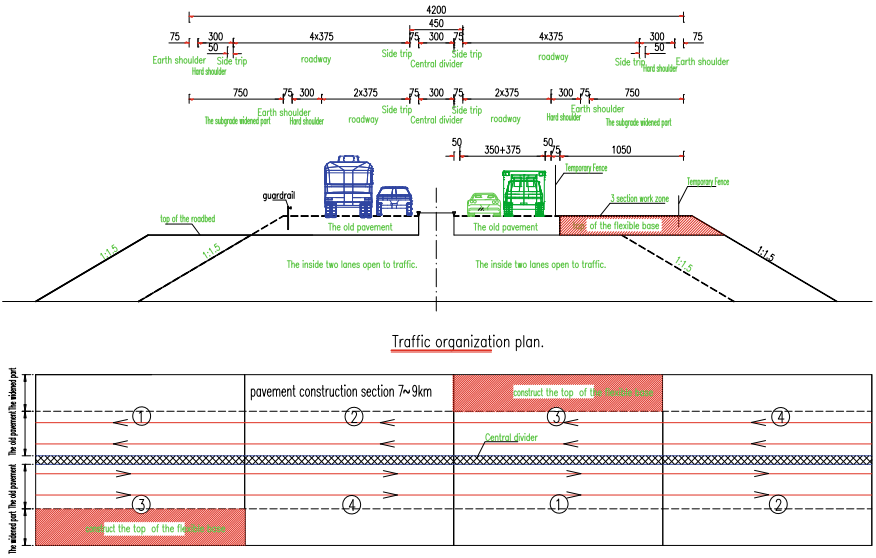


Fig. 3 Construction traffic organization (unit: cm)

Step 3 Transfer traffic, part 3 with two-way four-lane traffic, the part 1 overall paving road bed together first surface, and traffic safety facilities construction, setting the water horse temporary isolation between new and old pavement anti-collision facilities, about part 2, part 4 two-lane one-way traffic. In order to ensure the safety of driving, opens the mouth to consider using 40 km/h speed limit (Figs. 4 and 5).

Repeat the above steps separately until the construction is completed and the eight-lane cross section is formed (Table 2).

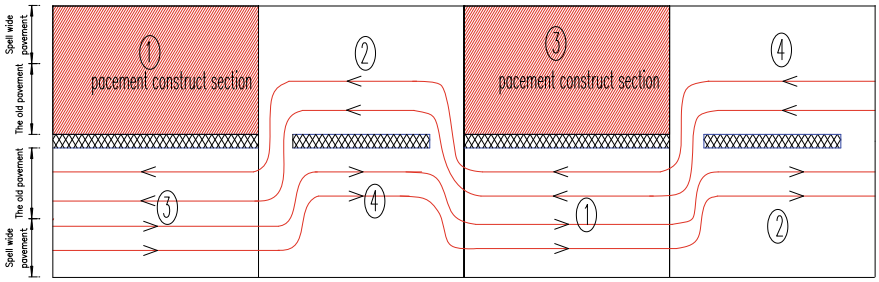


Fig. 4 Construction traffic organization in Step 3

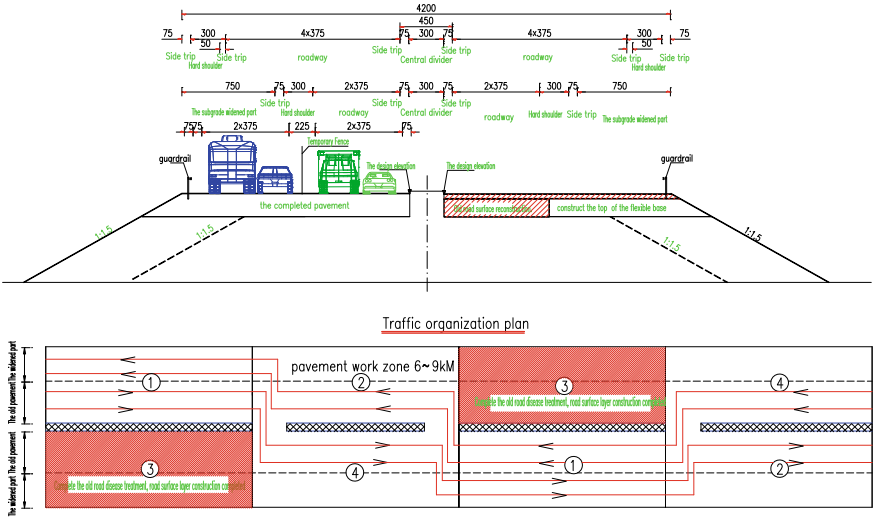


Fig. 5 Example of road surface construction traffic organization (unit: cm)

Table 2 Section of changzhang high-speed section division list

Section	The starting and ending pile number	Length (km)	Section number	Section start and stop pile number	Length (km)
Part 1	K29+500—k57+100	27.60	1	K39+587—k48+500	8.913
			2	K48+500—k57+100	8.60
Part 2	K57+100—k72+500	15.40	3	K57+100—k64+800	7.70
			4	K64+800—k72+500	7.70
Part 3	K72+500—k86+800	14.300	1	K72+500—k79+800	7.30
			2	K79+800—k86+800	7.00
Part 4	K86+800—k101+646.485	14.846	3	K86+800—k94+980	8.80
			4	K94+980—k101+646.485	6.67

3.2 Bridge Engineering

a. Corresponding subgrade construction, the first and second stage on both sides of subgrade in extension to road bed top stage and flexible base built on both sides of the road to the top surface at the same time, synchronous of construct a new wide bridge on both sides, set the temporary pavement and set up temporary isolation guardrail (Fig. 6).
b. Corresponding subgrade construction in the third stage, transfer the traffic to the sides, pavement inside path construct to the original road surface elevation at the same time, the Bridges synchronous transfer traffic to the side and maintain four lanes, Complete the reconstruction of the old bridge (Fig. 7).

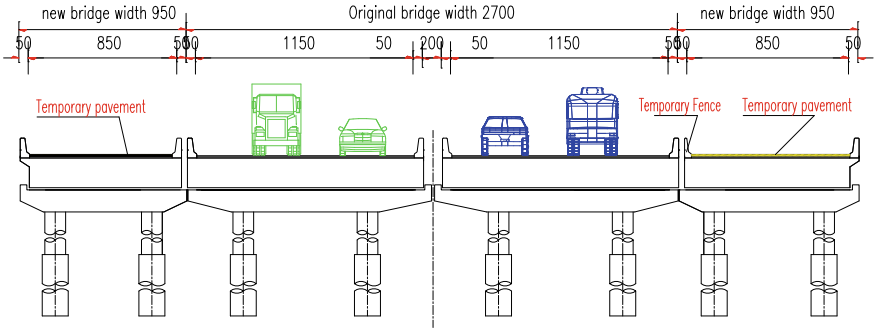


Fig. 6 Step 1 of the bridge construction

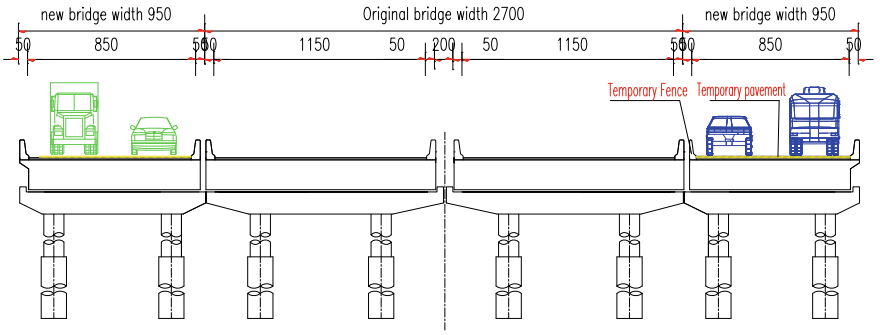


Fig. 7 Step 2 of the bridge construction

c. Corresponding subgrade construction in the fourth stage, pavement section unilateral construction surface at the same time, the Bridges synchronous to the introduction of the left (or right) traffic picture right (or left), joining together the left (or right) of old and new Bridges (Fig. 8).
d. In the fourth period of the construction of pavement, the section of the pavement is constructed on the single side of the road surface, and the traffic is converted to another one, and the construction step is repeated, and another new and old bridge is spliced (Fig. 9).
e. After the completion of splicing, the temporary guardrail was removed to complete the bridge construction.

3.3 Interconnect, Bridge and Other Special Points

In order to facilitate the transportation organization, the structure form of the upper span bridge is adopted in the same way as the continuous small box girder scheme

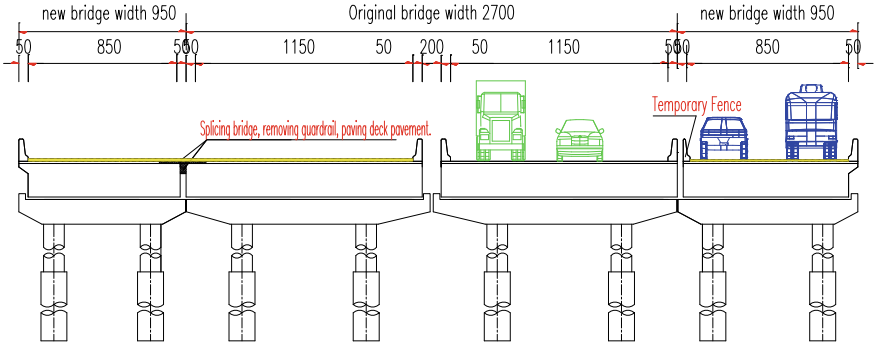


Fig. 8 Step 3 of the bridge construction

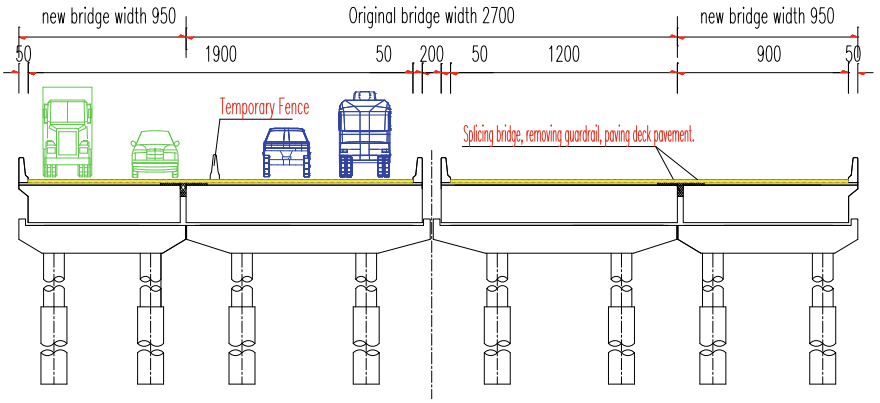


Fig. 9 Step 4 of the bridge construction

of the first simple support structure [2]. The bridge is uniformly concentrated in the old road and the traffic flow is concentrated during the blasting. The reform of intercommunication mainly adopts the principle of "first build and then disassemble", and the traffic flow in each direction should be maintained in all directions through the construction of the temporary road [3], which is limited to space and no longer needs to be discussed.

4 The Traffic Organization Implements Safeguards and Contingency Plans

Chang zhang expressway, while creating an overall traffic organization scheme of road traffic of traffic maintenance in the conversion of various construction areas, rescue, and equipped with a more detailed planning, and with the cycle of general

Table 3 The police force and emergency facilities (step 2)

Section	Section one		Section two		Section three		Section four	
Rescue		trailer		trailer		trailer		trailer
Police		assistant		assistant		assistant		assistant
Left	1（work zone）	2（Open traffic）	3（Open traffic）	4（Open traffic）	1（work zone）	2（Open traffic）	3（Open traffic）	4（Open traffic）
Starting	K39+587	K48+500	K57+100	K64+800	K72+500	K79+800	K86+800	K94+900
Destination	K48+500	K57+100	K64+800	K72+500	K79+800	K86+800	K941+900	destination
Length (Km)	8.913	8.6	7.7	7.7	7.3	7.0	8.18	6.667
Right	3（Open traffic）	4（Open traffic）	1（work zone）	2（Open traffic）	3（Open traffic）	4（Open traffic）	1（work zone）	2（Open traffic）
Police		assistant		assistant		assistant		assistant
Rescue		trailer		trailer		trailer		trailer
Section	Section one		Section two		Section three		Section four	
Note	The above police force is located in the corner of the non-construction work area. Each section is equipped with a patrol car and a crane, and 20 personnel and 1 traffic maintenance vehicle.							

traffic organization scheme section conversion real-time adjustment of police and facilities equipped with [3] (Table 3).

5 Experience in Traffic Organization and Some Thoughts

It is a new method of traffic organization to reconstruction and extension. Chang zhang expressway extension during the period of construction traffic organization scheme followed the project overall planning of traffic organization scheme, also combined with the actual situation in the process of actual implementation of the adjustment and optimization, in the process of implementation, accumulated a lot of valuable experience, there are some shortage, summarized as below:

(1) Section partitioning problem.

Section when the traffic organization design phase, the present situation of the zoning structure distribution was taken into account such factors as the position of the opening, workload balance between sections, but in the process of implementation, part of the section is still inadequate, such as part of the temporary passage 3, 4, section assigned to the uphill road, the road construction, caused the longer the vehicle queuing phenomenon, has formed a certain number of traffic jams.

(2) Reuse of road materials.

This scheme can realize the reuse of pavement materials through reasonable traffic conversion, It makes sense to save resources.

(3) Use a wide bridge to reuse

In order to ensure the security of the vehicle, the new bridge construct for an additional 2 m wider than pavement width. so, we set the guardrail, when opened to traffic after the completion of the project, it is used as the villagers along the river channel or change to the bridge maintenance and repair.

6 Conclusion

Around the new method of Traffic organization in the applications of Chang zhang expressway reconstruction and extension, fully considering the contractor with the operator, in ensuring safe operation at the same time, make full use of Chang zhang expressway resources, avoid the loss of tolls, both construction safety, efficiency, and effectively solves the problem of construction and operation. The practice verifies that this scheme is feasible.

With speed up the Midwest construction, the current domestic has large quantities of highway reconstruction, due to the low road network density in the Midwest, along the traffic during the reconstruction of road, the dependence of the reconstruction and construction period of maintaining the four-lane traffic organization and management of pioneering significance, the research results in jiangxi province and nationwide promotion, mature and can be provided for other similar highway reconstruction project experience for reference.

References

1. Zhang H (2015) With the “three guarantees” as the core concept of highway reconstruction and expansion of traffic organization management
2. Zhen Wu T (2006) The analysis and experimental study of the simple-continuous beam bridge. Master’s thesis of Zhejiang university
3. Zhang ZC (2015) The study of traffic organization of highway interchange. Master degree thesis of Changsha university of science and technology

The Research on Comprehensive Safety Analysis and Improvement Measures of Changzhang Expressway Yaohu Bridge and Houtian Terminal Interchange: Based on Lane-Changing Behavior

Chen Chen and Yazhen Z. Chen

Abstract The bridge has a strong closure itself and such as the factors that the bridge is high above the ground, which makes it difficult to rescue after the traffic accident. Because of the complexity of the traffic situation, the terminal interchange is the traffic accident prone area. The frequent incidents of expressway bridges and terminal interchange have severely affected the expressway safety, which has becomes one of the bottleneck of the freeway safety. Based on the characteristics of vehicle operation on the Yaohu bridge from Nanchang to Zhangshu expressway and Houtian terminal interchange, this study researched service level and the vehicle running status in weaving section, and then put forward the people-oriented of traffic management strategies and improvement measures which would significantly improve the safety level of large Bridges road and traffic efficiency.

Keywords Expressway · Weaving section · Traffic service level · Safety · Traffic management · Engineering improvement

1 The Introduction

With the rapid development of expressway transportation and the continuous increase speed of traffic flow, the number of traffic accidents is also increasing, which also brings huge hidden dangers to the safety of people's life and property [1]. It has been proved that that expressway accident is inextricably linked with the drive environment. The road section of Expressway Bridge usually has good Horizontal alignment and longitudinal linear, the speed of traffic flow is high, and the external environment

C. Chen (✉) · Y. Z. Chen
CCCC Second Expressway Consultant Co. Ltd., Wuhan, Hubei, China
e-mail: jg_cc@163.com

C. Chen
Traffic Engineering, Wuhan University of Technology, 498 Yingwu Road, Hanyang, Wuhan, Hubei, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_58

of the vehicle can easily cause traffic accidents. Hub of inward and outward expressway driving environmental change is the main line general section, and the inward and outward sections is a mixed area at the same time, and the off-ramp vehicles and main vehicles are jamming each other, which to form a traffic hazard [2]. The traffic conditions of the bridge and the interconnection is more complex, which is more likely to cause traffic accidents. In order to improve the safety of the combination road, it is necessary to study the vehicle behavior in the weaving section, and set up the corresponding traffic engineering facilities in a reasonable place, and strengthen the traffic management.

The expressway from Nanchang to Zhangshu (hereinafter referred to as the Changzhang expressway) is 85.65 km long, and it is the organic component of the national expressway network (7918), which is from Shanghai to kunming national expressway, which is also the major road segment of Jiangxi province's 'three-longitudinal and four-bar' road network main skeleton. It is the key point of the province of Jiangxi to connect the surrounding provinces to the provinces, to increase the external connection, it has very important position in the network. The traffic volume of Changzhang expressway reached more than 45,000 pcu/d, and the main traffic composition was heavy duty vehicle. There was an important section made up of Yaohu bridge and Houtian terminal interchange.

2 The Characteristics of the Vehicle Operation in the Intersections on the Yaohu Bridge to the Houtian Terminal Interchange

2.1 The Current Situation and Accident Analysis of the Junction of the Yaohu Bridge to Houtian Terminal Interchange

The Yaohu bridge is 9.1 km long, which is the longest expressway bridge and has the largest traffic volume in Jiangxi province. The average daily traffic volume is 21,646 (data of year 2010) before the expansion, with a growth rate of 19.1%. The traffic volume of this section is mainly truck. According to the traffic accident data from 2007 to 2011, the average number of accidents during the Yaohu bridge and the Houtian terminal interchange is less than the overall average (24.2%). The number of accidents is not prominent, but the degree of accident is quite serious and the casualty rate is relatively high.

Center of bridge pile No.ZK34 + 577.8, including pile No.K30-K34 has a higher accident rate (long straight sections), pile No.k38 has more accidents, the casualty rate was higher at k30-k32 and k38-k39, and the two places above were prone to accidents. The accident form is mainly followed by the collision, which is mainly

caused by "low visibility and not driving as required" and "with car too close" as (Figs. 1 and 2).

The specific accident road environmental cause analysis is as follows:

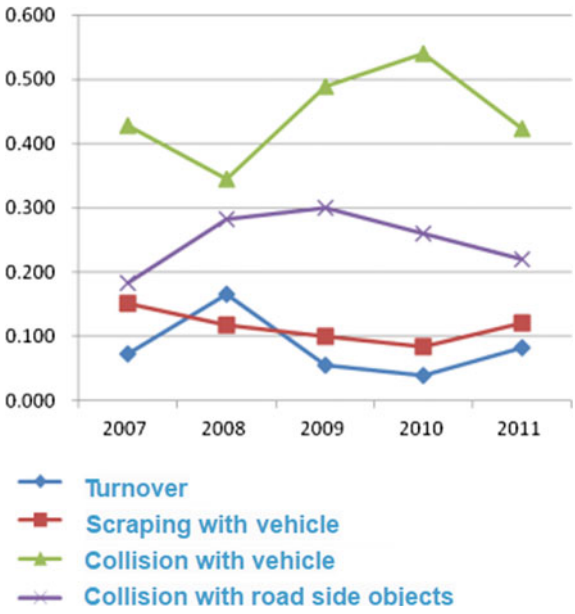


Fig. 1 Accident pattern

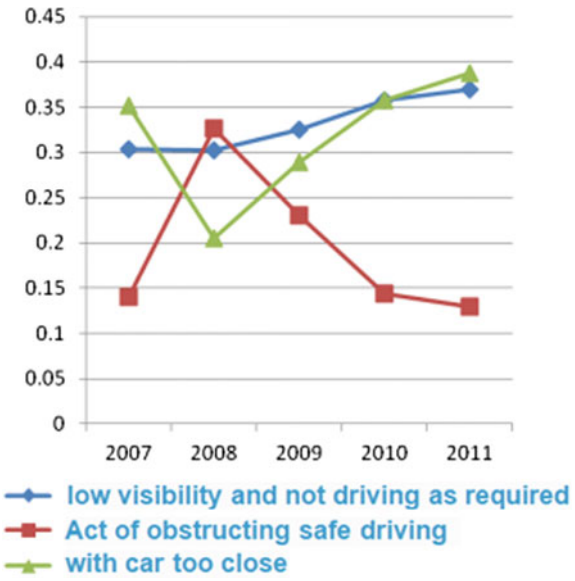


Fig. 2 Accident cause

Fig. 3 Scene photos of Yaohu bridge (2011)

(1) The Houtian terminal interchange near to the Yaohu bridge—is the houtian junction after the Yaohu bridge to NanChang, the vehicle has a high speed on the bridge section, vehicle lane changing is not timely to leave from expressway, or forced change leads to rear-end collision.
(2) The long straight road section—the starting point of the Yaohu bridge to Zhangshu includes a long straight line with a length of nearly 3 km, From the view of the safety, it is considered that the long straight section shall not be more than 20 times of the driving speed, with the calculation of design speed of 100 km/h, this long straight sections should not be greater than 2000 m. At present, the speed limit is 80 km/h, it is prone to fatigue due to that landscape of long straight segment. The driver is distracted and causes a traffic accident when he loses the right judgment (Fig. 3).
(3) The small longitudinal section—which is nearly 0% of the ground, The route design specification stipulates that the expressway synthetic longitudinal slope should not be less than 0.5%, especially on the bridge with a length of 9.1 km, the drainage pressure of the drainage hole is relatively large, and the surface water cannot be discharged in time, and the thick water film is formed on the road surface, which can threaten the safety of driving.
(4) Fog weather—The water area around this road section is developed, the humidity in the local environment is relatively large, and it is easy to fog on the bridge deck, and the high-speed vehicle suddenly enters the bridge. As the recognition is reduced, the vehicle entering the fog zone is prone to a sharp deceleration, followed by the braking of the vehicle, leading to the continuous rear-end.
(5) Landscape is monotonous—The guardrail and anti-dazzle panels on both sides of the viaduct formed a closed visual tunnel under the high speed train. Moreover, with long straight road sections, the drivers are prone to fatigue, lack of concentration, lack of speed, and can't help but accelerate, which tends to cause traffic accident in case of emergency.

2.2 The Expansion Plan of the Yaohu Bridge and the Traffic Organization of the Adjacent Section After the Expansion

Because the old Yaohu bridge bears little weight, the reconstruction and expansion plan is to separate the new two-way six-lane bridge from both sides, and the old bridge is reinforced and utilized (Fig. 4).

The huge bridge reconstruction and expansion of the huge Yaohu brigde has been adopted to separate the two sides of the original bridge. The upper structure is arranged as: $99 \times 20 + (4 \times 40 + 35) + (35 + 60 + 35) + 5 \times 35 + 3 \times 40 + 325 \times 20$ m, Total length 9106 m. The main use of $35 + 60 + 35$ m prestressed concrete cantilever casting box girder span Jinjiang VI level waterway. The approach bridge adopts a 20 m prestressed concrete split box girder, and the bridge hole corresponds to the original bridge (Fig. 5).

The section of the bridge head (near NanChang) of Changzhang high-speed YaoHu bridge is an integrated two-way 10 lanes, The YaoHu bridge is a separated two-way 10 lanes (an additional lane between the Houtian hubs), The head of YaoHu bridge (near ZhangShu) of the bridge is an integrated two-way 8 lane; The old bridge is used to drive small and medium-sized vehicles, and new Bridges on both sides are used to exercise heavy vehicles.

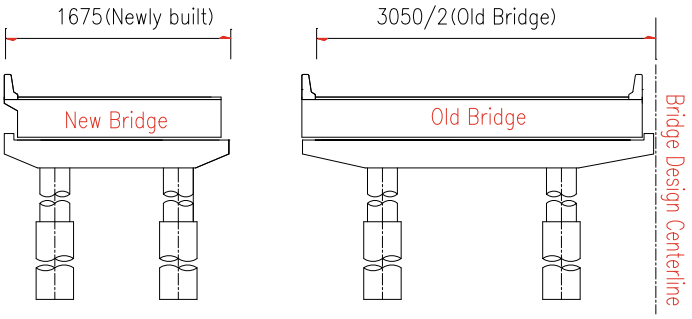


Fig. 4 Section view diagram of Yaohu bridge (cm)

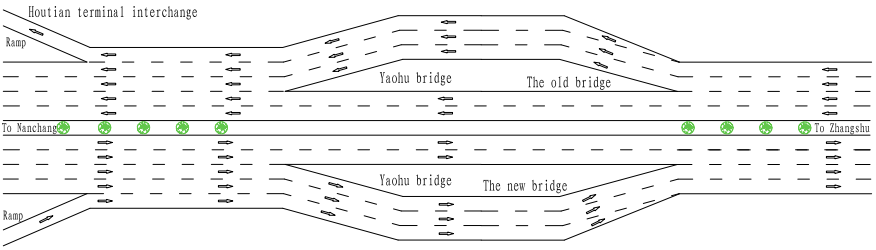


Fig. 5 Plane sketch of Yaohu bridge

2.3 Analysis of Traffic Lane Change in Mixed Area

Through the study on the lane change of the vehicle, the frequency of the potential lane change is increasing rapidly with the increase of the proportion of goods in the flow. The time and safety distance of the overtaking vehicle is determined by the vehicle speed and by the vehicle's body length of the super vehicle, so the flow ratio of the passenger truck and the flow ratio of the mini turcks and large cars are the main factors that determine the safety of the interlace area [3].

For expressway intersections, the security is mainly composed of the following four aspects:

(1) Intertexture length.

The length of the interlacing area is very short, which is equivalent to no signal control intersection. If the length of the intersect section is long, it's the equivalent of changing lanes on normal road. The mixed traffic flow is through the lane change to complete the interweaving operation, and the length of the intersections that limits the driver to complete all the required lane changes in a certain time and space. When the length of the interlaced section decreases, the frequency of lane change and the degree of turbulence generated will increase, and the security risk will also increase.

(2) Interwoven configuration [4]

Another important geometric feature of the interwoven region is the changing characteristics of the lane change in the interweaving area, which can be defined as the interweaving configuration, which mainly consists of three typical different interweaving configurations, namely the configuration A, B and C respectively (Fig. 6).

The number of lane change depends on the configuration of the intersecting road, and the interweaving configuration has a great influence on the utilization of the effective lane. For type A intersections, all vehicles must cross the arch line; In the B-type interweaving area, there is little restriction, and the interweaving operation produces traffic turbulence. Therefore, a hybrid car takes up more and more effective

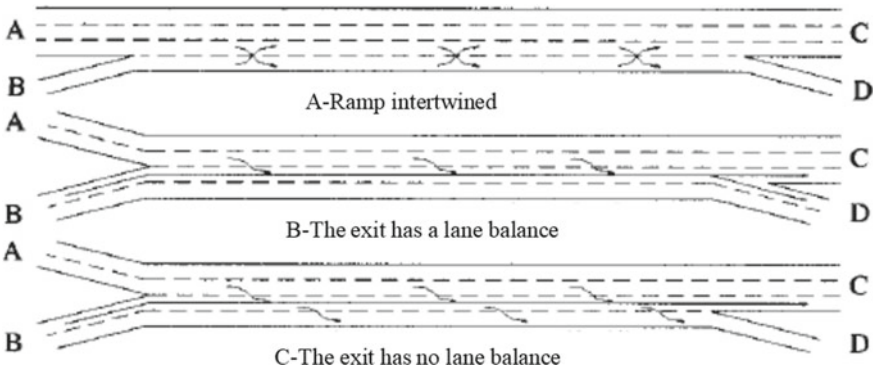


Fig. 6 Schematic diagram of interleaving area

space than a non-woven car. And the HouTian terminal interchange and YaoHu bridge section are type C interleaving area; however, if most vehicles choose the correct lane when entering the intersections, it will greatly reduce the interweaving of traffic.

(3) Interleaving flow rate

Mixed vehicles and non-woven vehicles are very different from the use of efficient lanes. So, the traffic rate of traffic flow in each direction is another important factor that affects the intersect section, including the total flow rate and the interweaving ratio. And the Interweaving ratio is the ratio of interwoven traffic in the interlace area.

(4) Design speed

It is similar to the confluence and cross through the gap. When a moving vehicle is ready to change lanes, it is necessary to judge the headway of the vehicle in the adjacent lane. The "critical space" as required for the vehicle to change lanes when the headstock of adjacent lanes is greater than the "critical space" required by changing lanes is a necessary condition for vehicle conversion lanes, and the main line of the interleaving area and the ramp speed are the most important factors.

2.4 Analysis of Service Level of Interwoven Area [5]

The main parameters of the service level of interweaving area is the traffic density, the code is K, and the formula is:

$$K = \frac{\left(\frac{Q}{N}\right)}{V} \tag{1}$$

The formula for the average velocity V for all the vehicles in the intersect section is as follows:

$$V = \frac{Q}{\left(\frac{Q_W}{V_W}\right) + \left(\frac{Q_{NW}}{V_{NW}}\right)} \tag{2}$$

The calculation formula of the average interwoven speed Vw and the average non-interwoven velocity Vnw is as follows:

$$V_i = 24 + \frac{V_{FF} - 16}{1 + W_i} (i = \text{w}/\text{nw}) \tag{3}$$

The calculation formula of the strength coefficient Wi is as follows:

$$W_i = \frac{a(1+Q_R)^b (Q/\mathrm{N})}{(3.28L)^d} \tag{4}$$

In which that: N is the total number of lanes in the interwoven segment; Q (pcu/h) is the total flow; V (km/h) is the average speed of all vehicles in the interwoven segment; Qw (pcu/h) is interwoven flow; Qnw (pcu/h) is non-interwoven flow; Vw (km/h) is the average interweaving speed of interwoven flow; VFF (km/h) is the average free flow velocity of the basic expressway sections, and the design speed is desirable too. Wi is the interweaving strength coefficient of mixed flow or unmixed flow; QR is the interwoven flow ratio; L(m) is the intersecting length; A, b, c and d are constants.

2.5 Operation Status Determination of Interwoven Zone

When Nw < Nwmax, it is unconstrained running state; When Nw is greater than or equal to Nwmax, it is constrained to run.

$$N_w = (1.21 \times N \times Q_R^{0.571} \times L^{0.234}) / V_w^{0.438} \tag{5}$$

In which that: Nwmax is the number of lanes that can occupy the most traffic for a given interweaving area; Nw is the number of lanes that the interweaving vehicle must occupy when it is unconstrained.

When calculating the Nw, first is to assume the constraint condition to determine the constants a, b, c, d value, using the formula (3) and (4) calculate the value of the Vw into (5) to calculate the value of the Nw, and then compared with Nwmax, after determining the constant value, according to the type (1)–(4), to calculate density K.

Design capacity Cd is according to the formula of expressway section design capacity in ⟨Design Specification for Highway Alignment⟩.

2.6 Evaluation of Changing Behavior Service Level

The road section from HoutIian terminal interchange to YaoHu bridge: the proportion of cars in this direction is only 29.9%, and the large and medium-sized vehicles is 70.1%. The cars don't need to change lanes and the coaches don't need to change lanes, which go straight, while the directly travelling vans (50 pcu/h) need to change lane for one time, and go to the newly-built bridge on the outside (as shown in Fig. 7 and 8); From the hub ramp to the main vehicle, trucks and buses do not need to change lanes. The cars (394 pcu/h, 11.3% of the total flow rate) needs to change into the inner lanes for three times, and the distance to be accommodated is about 872.857 m. According to the formula (1)–(5) and the calculation formula of the

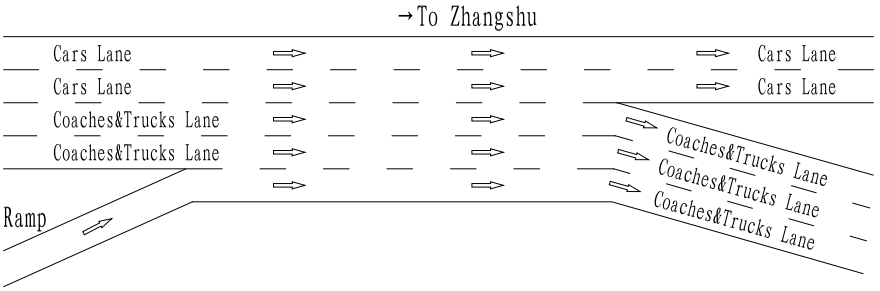


Fig. 7 Traffic organization diagram (to Zhangshu)

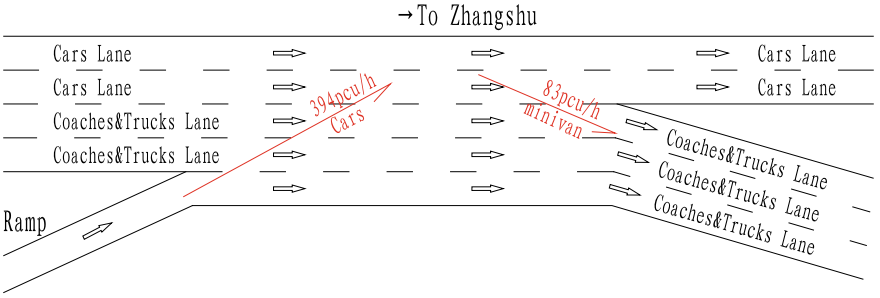


Fig. 8 Traffic lane change behavior diagram (to Zhangshu)

calculating capacity, the service level is about Level Two, which can fully meet the requirement of capacity and service level.

The road section from YaoHu to Houtian terminal interchange: The cars in this direction is 44%, and the large and medium-sized vehicles is 56%. The large vehicles on the right road need to change to the main lanes to turn left for one time (532 pcu/h), and the cars from Yaohu bridge need to turn right for 3 times to the right ramp (577 pcu/h), and the minivan (133 pcu/h) needs to change to the left lane for one time to go to the inside lane (as shown in Figs. 9 and 10). The acceptable lane

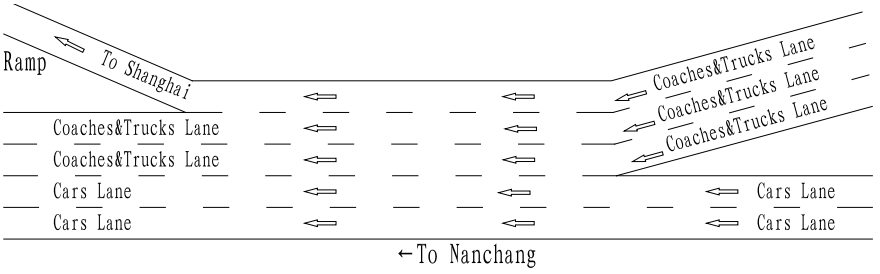


Fig. 9 Traffic organization diagram (to Nanchang)

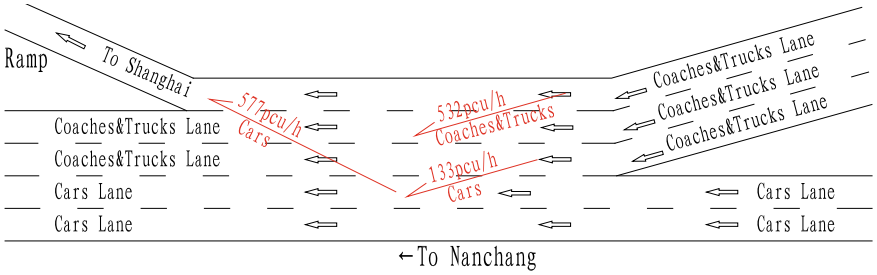


Fig. 10 Traffic lane change behavior diagram (to Nanchang)

change distance is approximately 995.351 m. According to the formula (1)–(5) and the calculation formula of the calculating capacity, the service level is about Level Three, which can fully meet the requirement of capacity and service level.

3 Traffic Management and Engineering Improvement Measures

Since the Yaohu bridge is separated bridge, it's not consistent with the subgrade cross-section before and after, the vehicles from the main lanes and the ramp would interflow with traffic flow of Nanchang to Zhangshu direction, and then divert to Yaohu old and new bridge (drive in the same direction separately). The vehicles on the main lanes from Zhangshu to Nanchang should be separated to the Yaohu by the vehicle model and then interflowed, and then separated from the exit ramp at last. The traffic organization is complex, and the speed is higher than other sections due to the long straight road section. While the traffic capacity and service level meet the requirements, but on the basis of safety consideration, in the premise of enhancing the water drainage on the road, it is proposed to reduce the risk of accidents of interleaving road section by improving traffic guiding facilities, strengthening passive protection and traffic management.

3.1 Traffic Guidance Facilities

The road traffic guidance facilities are set up on the section of the Yaohu bridge to inform the drivers of the driving environment in advance and to induce the vehicle to adopt a rational running path.

Fig. 11 Traffic guidance system diagram

Fig. 12 The speed limit signs

(1) To set up the driving and forward shunt driving warning signs to make different vehicle types in different lanes at 2 km, 1.6 km and 500 m in front of the separation roadbed of the great YaoHu bridge [6]. At the same time, it's necessary to set the ground markings to match the signs (Fig. 11).

Setting the "shunt driving" warning signs aims to let them drive freely, so as to reduce the number of vehicles changing lanes and reduce the traffic accidents caused by the traffic confluence. To set a multi-level "vehicles interring the main lanes to the right" sign, it is helpful to prompt the vehicles from ramp on the main lanes to drive on the right lane, and reduce the number of traffic confluence.

(2) To set the speed limiting sign at the starting of the Yaohu bridge. Considering the high accident rate of rain and fog in this section, the old and new bridge should be provide with a limiting sign according to the climate and vehicle model (Fig. 12).
(3) To prevent speeding driving due to lack of change of road environment. The longitudinal deceleration lines are set at the edge of the long straight road segment lanes [7].
(4) To set up the red speed bump. Accident rate is higher at the Yaohu bridge, the red speed bump setting on both ends of the bridge and in the middle of the bridge will help increase the coefficient of friction of the road, and warn the drivers to slow down [8].
(5) The yellow flashing warning lights are set up at the starting point of the separated roadbed and the entrance ramp, and the warning columns are set in the location of the zebra crossing, which is helpful to highlight the exit of the ramp.
(6) Set the bridge name signs on both starting points of the bridge, and it is suggested that "the length of the bridge is 9.1 km". It is helpful for drivers to prepare for long time driving on the bridge.

3.2 Passive Protection Facilities

The speed on main lanes is high, bumper is set up at the starting point of the separated roadbed which can be used to guide the exit ramp and also can help to reduce the damage caused by the collision end of the vehicle.

The bridge is divided into two sections with a length of 60 m medial strip opening, and a movable guardrail is set up at medial strip opening. In case of emergency, it can be turned on and used to evacuate traffic to the opposite lanes.

3.3 Traffic Management Means

(1) Electronic police and speed limiting sign are set up at the starting points of Yaohu bridge and Houtian terminal interchange.
(2) Sign with "Rain and Foggy Weather, Slow Down" is set every 1 km on Yaohu bridge.
(3) The traffic accident of the Yaohu Bridge is mostly concentrated in the evening, and it is easy to foggy on the bridge. Then it is very necessary to set up yellow fog lights, active luminous raised road signs to strengthen the lighting on the bridge [9]. Considering that the traffic from the Houtian terminal interchange to the Yaohu bridge section is complex and the lane changing behavior is common, it is recommended that the Houtian terminal interchange-Yaohu bridge section (two-directions) be provided with lighting to reduce the probability of an accident occurring at night when the vehicle changes lanes.

4 Conclusions

Based on the characteristics of vehicle operation in the interweaved area, the traffic guidance facilities of the Changzhang Expressway Yaohu Bridge and the Houtian terminal interchange section are optimized, and passive protection and traffic management methods are strengthened. As a result, the vehicle routing and driving strategies in the interlaced area of the section are optimized, and the Changzhang Expressway Yaohu Bridge and the Houtian terminal interchange section are highly praised by the management units and users in the later stage of operation. The number of accidents on this section has been significantly reduced. It has been proved by engineering practice that the crossover model in weaving section is feasible. It also provides an example for safe operation of the similar road section.

References

1. Wang H (2009) Present situation and countermeasure of traffic safety in high-grade highway. J Saf Sci Technol. June
2. Chen Q (2015) Study on the safeguard technology of Naqian expressway interlaced area. J Highw Transp Res Dev
3. Jiang R (2005) Studies on the safety of basic freeway session. College Station
4. Transportation Research Board. Highway capacity manual
5. Design Specification for Highway Alignment (JTG D20-2017). Ministry of Transport of the People's Republic of China. January, 2018
6. Zhou Y (2011) Discussion on the dividing of the old road widening on one side of freeway. J China Foreign Highw
7. Zhang K, Qiu Q (2014) Study on deceleration effect of visual speed reduction markings on the road. J Saf Sci Technol
8. Zhang B (2011) Driver factors and freeway traffic safety. Highway & Transpotation in Inner Monglia
9. Ma H, Yang R, Reng Z, Zhang Z (2006) Study and application on safety pre-assessment of freeway construction project. J Saf Sci Technol

Study on the Stability Control Strategy for Distributed Driving Electric Vehicle

Deng Hai, Xianyi Xie and Lisheng Jin

Abstract This paper describes a stability control strategy for four wheels distributed driving electric vehicle (4WD). The stability algorithm consist of two parts: (1) The upper hierarchical controller that determine the required yaw moment which used to maintain the vehicle stability; (2) The lower hierarchical controller that distribute the driving or braking force of each wheels. The upper hierarchical controller calculate the yaw moment through sliding model theory which take account into the error between actual value and idealized value for yaw rate and side slip angle. The lower hierarchical controller distribute the required yaw moment by adjusting braking or driving torque based on tire load condition. Firstly, the required yaw moment is preferred to distribute by the hub wheel motor. Then the hydraulic brake system compensates for the required yaw torque, when the hub motor outputs the maximum torque. Numerical simulation studies have been conducted to evaluate the proposed vehicle stability control strategy. The simulation results demonstrated that the proposed stability control strategy can improve the vehicle handling stability and driving safety.

Keywords Distributed driving electric vehicle · Yaw moment · Sliding model control · Quadratic programming · Handling stability

Corresponding author has the same contribution as the first author.

D. Hai
Integration Platforms Department, National Engineering Research Center of Railway Vehicles, Changchun, China
e-mail: denghai@cccar.com.cn

X. Xie (✉)
School of Mechanical & Electrical Engineering, Changchun Institute of Technology, Changchun, China
e-mail: xiexianyi123@126.com

L. Jin
Transportation College, Jilin University, Changchun 130024, China
e-mail: jinls@jlu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_59

1 Introduction

Vehicle stability control system refers to the vehicle control device through real-time adjustment of vehicle operating status. It is also an active safety device to avoid vehicle instability so that the vehicle can follow the driver's expectations [1].

Vehicle stability control system has been developed for decades, such as ESP (electric stability program) [2–4], 4WS (four wheel steering system) [3, 4], AS (active steering, include AFS and ARS) [2, 4–6] has been more and more applied to passenger vehicle. The advantage of active steering is that improves handling stability of the vehicle in a linear state. While the vehicle in a non-linear state, the effect is significantly reduced. So vehicle stability control system by controlling longitudinal force has become important active safety equipment of the vehicle. With the increasing demand for vehicle handling performance, different control strategies have been emerging. There are some scholars [7, 8] combine the active steering technology and direct yaw moment (DYC) control system through some strategy, although the simple controllers can fully utilize of itself and reduces the degree of speed decline, but also the active steering system needs to be added to some mechanical devices, which makes the chassis system more complex. There are also many scholars [9] based on the β phase plane to determine the stability of the vehicle, reducing the intervention frequency of stability controller, but still cannot change the facts of speed down, while it also need to install vehicle sensor to estimate the side slip angle. The chassis structure of the conventional internal combustion engine vehicles only rely on the way of differential braking or active steering to produce yaw moment to maintain a stable driving. While if a large yaw moment is required, more tire utilization and energy are necessarily consumed. Reduce the stability margin, will reduce the vehicle speed, have an impact on the longitudinal motion of the vehicle. Especially, in some cases, the vehicle did not out of control. But with the intervention of the stability control system, reducing the handling stability of vehicles. Also these stability control system cannot take into account the driver's driving intention.

However, the advantages of distributed driving electric vehicles has been more and more obvious, torque output response of the motor is faster and more efficient. The power trains of these types of vehicles consist of two or four electric motors, which are integrated into each drive wheel and can be controlled independently [5–8]. The control of the vehicle dynamic through proportioning of the vehicle tractive force is easily accessible. It also provides more diversified solution for vehicle dynamics control. The braking torque acts on the vehicle can not only be provided by the hub motor, but also rely on the hydraulic system to provide [10].

Compared to hydraulic braking system, the braking torque provided by the motor is limited. While the peak torque of wheel motor at high speed cannot meet the stability of its torque control needs [11]. Therefore, in order to guarantee the safety of vehicle driving, distributed electric vehicle chassis structure still retains the hydraulic brake system. When the motor max torque cannot meet the required yaw moment, then rely on hydraulic system compensation.

Depend on the author's previous research, the fuzzy control theory is used to design the vehicle stability controller [9]. The design process of the fuzzy controller requires plenty of practical engineering experience and parameter debugging. Although the genetic algorithm is used to optimize the parameters and membership function of the fuzzy controller, controller performance has improved, but it still limited. Due to its robustness, sliding mode control (SMC) has increasingly been applied in vehicle traction control [12–16]. This paper chooses sliding model control theory which no need for accurate modeling and easy to adjust parameter [7], used to calculate the yaw moment which to maintain the vehicle stability.

In the case of vehicle instability, it is usually caused by the tire force saturation phenomenon [12, 13], cannot provide sufficient lateral force, resulting in side slip or instability. The tire longitudinal force utilization can characterize the stability margin of the tire [14, 15]. Therefore, the design of stability controller, should consider such a problem. Although many scholars [15, 16] designed the controller based on the control of the tire slip rate, the principle is similar. When the tire slip rate is close to 100%, the tire force also will be saturated. Therefore, the control of tire utilization also indirectly controls the slip ratio of the wheels.

This paper takes the minimum utilization of all tires as the optimization target, the yaw moment distributor is designed according to the quadratic programming theory. Considering the limit of the motor output torque and the braking system. And a first-order inertia is added to characterize the delayed response of these system.

This paper is organized as follows. In Sect. 2, presents the additional yaw moment which maintain the vehicle stability by sliding model controller and torque distribution controller, respectively. Computer simulation results of the vehicle with driver model tracking a DLC test are presented in Sect. 3. Finally, Sect. 4 gives a brief conclusion.

2 Vehicle Stability Controller

The controller uses the yaw rate and side slip angle as inputs in order to compute the total yaw moment and lateral force by the sliding model control (SMC) theory.

The sliding model control strategy has several advantages, such as easy to adjust parameter and no need for precise modeling. And many scholars define the sliding model surface as follow [14, 15].

$$S = w_r - w_{rd} + fai(\beta - \beta_d) \tag{1}$$

$$\dot{S} = \dot{w_r} - \dot{w_{rd}} + fai\left(\dot{\beta} - \dot{\beta_d}\right) \tag{2}$$

Obtained from equation as follow:

$$I_Z \cdot \dot{w}_r - I_{XZ} \cdot \ddot{\varphi} = a\big(F_{yfl} + F_{yfr}\big)\cos\delta + \frac{d}{2}\big(F_{yfl} - F_{yfr}\big)\sin\delta - b\big(F_{yrl} + F_{yrr}\big) + \frac{d}{2}\big(F_{xfr} - F_{xfl}\big)\cos\delta + a\big(F_{xfl} + F_{xfr}\big)\sin\delta + \frac{d}{2}(F_{xrr} - F_{xrl}) \tag{3}$$

Due to the value of δ is very small, so take the $\sin\delta = 0$. In order to reduce the complexity of the controller. Therefore, the above formula can be simplified as follows:

$$I_Z \cdot \dot{w}_r = a\big(F_{yfl} + F_{yfr}\big)\cos\delta - b\big(F_{yrl} + F_{yrr}\big) + \frac{d}{2}\big(F_{xfr} - F_{xfl}\big)\cos\delta + \frac{d}{2}(F_{xrr} - F_{xrl})$$

Command the follow formula:

$$M_z = \frac{d}{2}\big(F_{xfr} - F_{xfl}\big)\cos\delta + \frac{d}{2}(F_{xrr} - F_{xrl}) \tag{4}$$

The formula (4) present the yaw moment to be applied to maintain vehicle stability.

As above discuss, if the vehicle driving on the low adhesion road, maintain the sideslip angle in a reasonable range to maintain vehicle driving stability has great significance. For the high adhesion road, tracking the ideal yaw rate is more to enhance vehicle handling stability significance. But the above controller only based on the *fai* to determine the control law of the sliding model. In order to obtain the sliding control law accurately. Therefore, the sliding control law should be changed by side slip angle of the vehicle. So design the sliding model control law as follow.

2.1 Sliding Model Variable Structure Control of Yaw Rate **Wr-*SMC***

Set the error of the yaw rate as:

$$e = w_r - w_{rd} \tag{5}$$

Define the sliding model surface as:

$$S = \dot{e} + e \tag{6}$$

While $s \to 0$, so $e \to 0$, $\dot{e} \to 0$, and then $S = \dot{e} + e$

$$\begin{aligned} S &= \dot{e} + e = (\dot{w}_r - \dot{w_{rd}}) + w_r - w_{rd} \\ &= \frac{2(ak_1 - bk_2)\beta}{I_Z} + \left[\frac{2\left(a^2k_1 + b^2k_2\right)}{v_x \cdot I_Z} + 1\right] w_r - \frac{2ak_1\delta_f}{I_Z} - \dot{w_{rd}} - w_{rd} \\ &= \frac{2(ak_1 - bk_2)\beta}{I_Z} + \left[\frac{2\left(a^2k_1 + b^2k_2\right)}{v_x \cdot I_Z} + 1\right] w_r - \frac{2ak_1\delta_f}{I_Z} - \dot{w_{rd}} - w_{rd} + \frac{\Delta M_r}{I_Z} \end{aligned} \tag{7}$$

Derivation of left and right side of the equation:

$$\dot{S} = \ddot{e} + \dot{e} = (\ddot{w}_r - \ddot{w_{rd}}) + (\dot{w}_r - \dot{w_{rd}}) = A + \frac{\Delta \dot{M}_r}{I_Z} \tag{8}$$

wherein,

$$\begin{aligned} A = \frac{2(ak_1 - bk_2)}{I_z}&\left[\frac{2(a+b)}{mv_x}\beta + \left(\frac{2(ak_1 - bk_2)}{mv_x^2} - 1\right) w_r - \frac{2a}{mv_x}\delta_f\right] \\ &+ \left(\frac{2\left(a^2k_1 + b^2k_2\right)}{v_x \cdot I_Z} + 1\right)\dot{w}_r - \frac{2ak_1}{I_Z}\dot{\delta_f} - \dot{w_{rd}} - \ddot{w_{rd}} \end{aligned} \tag{9}$$

By using the law of reaching law at constant speed to design the control law for sliding model.

We set:

$$\dot{S} = -K_r \cdot sgn(s),\ K_r > 0, \tag{10}$$

And then,

$$\Delta \dot{M}_r = -I_Z[K_r \cdot sgn(s) + A] \tag{11}$$

After integral the $\Delta \dot{M}_r$, ΔMr is obtained which adopted to maintain the vehicle driving stability.

2.2 *Sliding Model Variable Structure Control of Side Slip Angle β-SMC*

Set the error of the side slip angle as:

$$e = \beta - \beta_d \tag{12}$$

While $s \to 0$, so $e \to 0$, $\dot{e} \to 0$, and then $S = \dot{e} + e$

$$S = \dot{e} + e = \left(\dot{\beta} - \dot{\beta}_d\right) + \left(\beta - \beta_d\right) = \left(\frac{2(k_1 + k_2)}{mv_x} + 1\right)\beta + \left(\frac{2(ak_1 + bk_2)}{mv_x^2} - 1\right) w_r - \frac{2k_1}{mv_x}\delta_f - \dot{\beta}_d - \beta_d \tag{13}$$

Derivation of left and right side of the equation:

$$\dot{s} = \ddot{e} + \dot{e} = \left(\frac{2(k_1 + k_2)}{mv_x} + 1\right)\dot{\beta} + \left(\frac{2(ak_1 + bk_2)}{mv_x^2} - 1\right)\dot{w}_r - \frac{2k_1}{mv_x}\dot{\delta}_f - \ddot{\beta}_d - \dot{\beta}_d = B + \left(\frac{2(ak_1 + bk_2)}{mv_x^2} - 1\right)\frac{\Delta M_\beta}{I_Z} \tag{14}$$

By using the law of reaching law at constant speed to design the control law for sliding model.

We set: $\dot{S} = -K_\beta \cdot sgn(s)$, $K_\beta > 0$, and then

$$\Delta M_\beta = I_Z\left[B + K_\beta \cdot sgn(s)\right]\left(1 - \frac{2(ak_1 + bk_2)}{mv_x^2}\right)^{-1} \tag{15}$$

Inside,

$$B = \left(\frac{2(k_1 + k_2)}{mv_x} + 1\right)\left[\frac{2(k_1 + k_2)}{mv_x}\beta + \left(\frac{2(ak_1 + bk_2)}{mv_x^2} - 1\right)w_r - \frac{2k_1}{mv_x}\delta_f\right] + \left(\frac{2(ak_1 + bk_2)}{mv_x^2} - 1\right)\left[\frac{2(ak_1 - bk_2)}{I_Z}\beta + \left(\frac{2\left(a^2k_1 + b^2k_2\right)}{v_x I_z}\right)w_r - \frac{2ak_1}{I_Z}\delta_f\right] - \frac{2k_1}{mv_x}\dot{\delta}_f - \ddot{\beta}_d - \dot{\beta}_d \tag{16}$$

2.3 Test the Stability of Control Law

The second stability criterion of Lyapunov is used to prove the stability of the controller, set the Lyapunov function as, $V_L = S^2/2$, and then:

For the *Wr*-SMC has:

$$\dot{V}_L = s \cdot \dot{s} = s\left(A + \frac{\Delta\dot{M}}{I_Z}\right) \tag{17}$$

Take this into the expression of $\Delta\dot{M}_r$

$$\dot{V}_L = s \cdot \dot{s} = s\left[\mathrm{A} - (k_r \cdot \mathrm{sgn(s)} + \mathrm{A})\right] = \mathrm{s}\left[-K_r \cdot \mathrm{sgn(s)}\right] = -K_r|s| \le 0 \tag{18}$$

Similarly, for the *Wr*-SMC has:

$$\dot{V}_L = s \cdot \dot{s} = s\left[\mathrm{B} - \left(k_\beta \cdot \mathrm{sgn(s)} + \mathrm{B}\right)\right] = \mathrm{s}\left[-K_\beta \cdot \mathrm{sgn(s)}\right] = -K_\beta |s| \leq 0 \quad (19)$$

2.4 *Coordinated Control Mechanism for the* Wr-*SMC and* β-*SMC*

Depend on the $|\beta|$ value, adjusting the weighting coefficients of ΔMr and ΔM_β. K is determined according to the relationship between vehicle stability and $|\beta|$:

$$K = \begin{cases} 1, |\beta| \leq \beta_0 \\ 1 - \frac{|\beta| - \beta_0}{\beta_1 - \beta_0} \\ 0, |\beta| \geq \beta_1 \end{cases}, \ \beta_0 < |\beta| < \beta_1 \quad (20)$$

So, the yaw moment acting on the vehicle is:

$$\Delta M = K \cdot \Delta M_r + (1 - K) \cdot \Delta M_\beta \quad (21)$$

However, as described in the "Introduction" section, it is concluded that if the yaw moment is assigned in accordance with the conventional differential braking form, it will cause the undesirable decrease in the total speed of the vehicle and lead to negative effect on comfortable for passenger. Under the extreme conditions, once the tire force was saturated, the vehicle is prone to slip caused instability phenomenon. Thus, this paper depend on tire utilization of each wheel to design the yaw moment distributor based on the quadratic programming theory. Designing yaw moment distributor based on vehicle load.

3 Double Lane-Change Test

In order to verify the validity of the proposed stability control strategy, the double-lane-change test simulation were carried out. During the simulation test, the initial speed is 120 km/h, road adhesion coefficient is 0.8. Without active accelerate or decelerate input from the driver. As a comparison baseline, a widely referred stability control strategy based on differential braking control which discussed in [9] and the "uncontrolled vehicle" (FWS) all add to test together. The test results are shown in Fig. 1.

In the simulation, the "uncontrolled vehicle" ("FWS") failed to complete the double lane-change test as shown in Fig. 1a, "Stability control based on differential

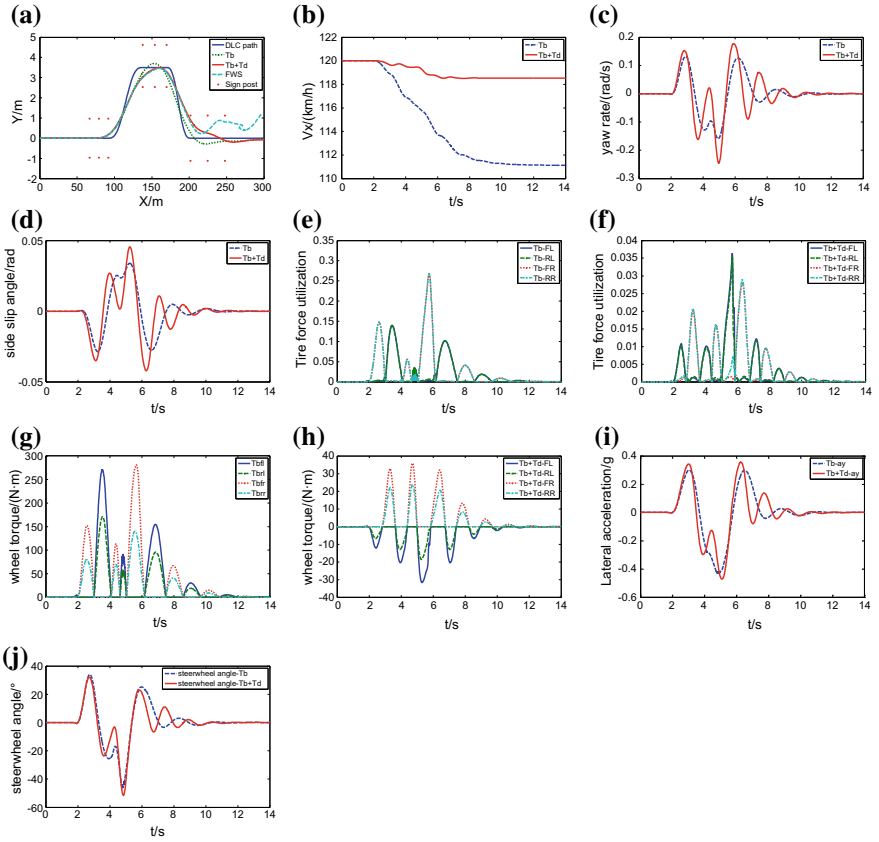


Fig. 1 Simulation results

braking" ("Tb") and "Stability control based on quadratic programming" ("Tb + Td") all completed double-lane-change test, with no collision accident or occur instability.

Figure 1c, d and (i) demonstrated the yaw rate, slip angle, lateral acceleration response under different control strategies ("Tb" and "Tb + Td"), respectively. The response of yaw rate and lateral acceleration is reasonable, the side slip angle is in ideal range (no more than 0.14 rad). As in Fig. 1b, the velocity of the vehicle with "Tb" has a greater degree of decrease than "Tb + Td". The reason is that "Tb" produce yaw moment only rely on applying differential brake force to the wheels, it will lead to a decrease in vehicle speed inevitably. In order to obtain the same size of the yaw moment, the "Tb" consume more braking torque to keep the vehicle driving safety (in Fig. 1g, h). This also leads to the increased utilization of the longitudinal force of the tire for "Tb" directly (in Fig. 1e, f).

On the contrary, for "Tb + Td", the same size of the yaw moment is generated by the appropriate driving and braking force, and reduce the number of interventions

by stability controller, the reduction of decline of speed and the tire utilization rate is also decreased.

4 Conclusion

In this paper, the eight-degree-of-freedom vehicle dynamics model, the driver model and the ideal model are established. Using the hierarchical design method, the yaw moment calculator is designed by the sliding model control theory, and the torque distribution is designed by tire load condition. Priority is based on the hub motor assignment, if the required maximum torque exceeds the wheel motor limit, then start the hydraulic brake to compensate. The closed-loop double line change simulation test was carried on the Matlab/simulink platform. The aforementioned objectives have been achieved by using the proposed control system. The simulation results demonstrated that the vehicle trajectory accords with the driver's intention and ensure driving safety, not only reduce the impact on the speed of the vehicle, but also decrease the tire force utilization and intervention numbers of stability control system.

Acknowledgements This work is supported by the National Natural Science Foundation (no. 51575229), the National key R & D project of China (no. 2017YFB0102600), and the Electric intelligent vehicle innovation team of the Science and Technology Department of Jilin Province (no. 20180519022JH).

References

1. Nakazawa M, Isobe O, Takahashi S et al (1995) Braking force distribution control for improved vehicle dynamics and brake performance. Veh Syst Dyn 24(4–5):413–426
2. He J et al (2006) Coordination of active steering, driveline, and braking for integrated vehicle dynamics control. Proc Inst Mech Eng, Part D: J Automob Eng 220(10):1401–1420
3. He J (2005) Integrated vehicle dynamics control using active steering, driveline and braking. Dissertation, University of Leeds
4. Tavasoli A, Naraghi M, Shakeri H (2012) Optimized coordination of brakes and active steering for a 4WS passenger car. ISA Trans 51(5):573–583
5. Yin G, Wang S, Jin X (2013) Optimal slip ratio based fuzzy control of acceleration slip regulation for four-wheel independent driving electric vehicles. Math Probl Eng 2013
6. Wong JY (2001) Theory of ground vehicles, 3rd edn. Wiley, New York
7. Mousavinejad E, Han QL, Yang F et al (2017) Integrated control of ground vehicles dynamics via advanced terminal sliding mode control. Veh Syst Dyn 55(2):268–294
8. Yim S, Kim S, Yun H (2016) Coordinated control with electronic stability control and active front steering using the optimum yaw moment distribution under a lateral force constraint on the active front steering. Proc Inst Mech Eng, Part D: J Automob Eng 230(5):581–592
9. Jin L, Xie X, Shen C et al (2017) Study on electronic stability program control strategy based on the fuzzy logical and genetic optimization method. Adv Mech Eng 9(5):1687814017699351
10. Chen Y, Wang J (2012) Design and evaluation on electric differentials for overactuated electric ground vehicles with four independent in-wheel motors. IEEE Trans Veh Technol 61(4):1534–1542

11. Liu H, Chen X, Wang X (2012) Overview and prospects on distributed drive electric vehicles and its energy saving strategy. Prz Elektrotech 88:122–125
12. Amodeo M, Ferrara A, Terzaghi R et al (2010) Wheel slip control via second-order sliding-mode generation. IEEE Trans Intell Transp Syst 11(1):122–131
13. Cho K, Kim J, Choi S (2012) The integrated vehicle longitudinal control system for ABS and TCS In: IEEE international conference on control applications (CCA). IEEE, pp 1322–1327
14. Tanelli M, Ferrara A (2013) Wheel slip control of road vehicles via switched second order sliding modes. Int J Veh Des 62(2–4):231–253
15. Drakunov S, Ozguner U, Dix P et al (1995) ABS control using optimum search via sliding modes. IEEE Trans Control Syst Technol 3(1):79–85
16. Unsal C, Kachroo P (1999) Sliding mode measurement feedback control for antilock braking systems. IEEE Trans Control Syst Technol 7(2):271–281

Regression Tree Model of the Scale's Dynamic Adjustment of Cruising Taxicab Capacity

Xiaofei Ye, Min Li, Xuan Li, Lili Lu and Yu-ming Jin

Abstract Based on the taxi operation datasets from Taxi Information Management System in Ningbo City, a regression tree model for the scale's dynamic adjustment of cruising taxi capacity was established by considering the average daily loading time of single taxi, mileage utilization and other indexes. The scale's dynamic adjustment mechanism of cruising taxi capacity and thresholds standard of key indicators were proposed by coupling the functions on the balance between supply and demand and the importance ordering relationship of indicators, which consisted of the taxi ownership per person, the sharing ratio of taxi in public transport trip structure, and other indexes. The results indicate that, (1) in the three-layer structure of regression tree; mileage utilization has the strongest effect on the scale of cruising taxicab capacity. Then the average daily revenue of single taxi, the average waiting time of single taxi's carrying, the average operation time of single taxi, and the revenue per 100 km have the stronger influence which decreased progressively. And the average daily loading time of single taxi is chosen as the third layer of classification indicator; *p*-values of the indicators in every layer are less than 0.05 and all passed the significance tests. The standard error and the ratio of mis-discrimination between training and testing samples are 6.13% and 0.07 which indicate the overall accuracy of model is better. (2) When mileage utilization is less than 0.6179 and the average daily revenue of single taxi is less than 798.38 Yuan, the scale of cruising taxi capacity is surplus and need to be reduced. (3) When mileage utilization is more than 0.6774 and the average waiting time of single taxi's loading is more than 259.09 s, the scale of cruising taxi capacity is insufficient and suggested to increase 463 taxies.

X. Ye (✉) · M. Li · X. Li · L. Lu
School of Maritime and Transportation, Ningbo University, Ningbo 315211, China
e-mail: yexiaofei@nbu.edu.cn

X. Ye · X. Li · L. Lu
Ningbo Port Trade Cooperation and Development Collaborative Innovation Center, Ningbo 315211, China

Ningbo University Sub-Center, National Traffic Management Engineering and Technology Research Center, Ningbo 315211, China

Y. Jin
School of Economics and Management, Chang'an University, Chang'an 710021, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_60

Keywords Traffic engineering · The scale of capacity · Regression tree · Cruising taxi · Threshold analysis

1 Introduction

The taxi is an important part of the comprehensive transportation system of the city [1], but the taxi industry has been greatly impacted by the development of the ride-hailing taxi [2, 3]. For this reason, the Ministry of Transport in China announced that it is necessary to establish a reasonable scale dynamic monitoring and adjustment mechanism for taxi capacity to gradually achieve market regulation [4].

Domestic and foreign scholars carried out in-depth studies on the scale forecasting and monitoring evaluation of taxi capacity [5–8], which pointed out that the key indexes and thresholds of the taxi capacity scale adjustment include: the taxi ownership per person (between 20 and 30 cars per 10,000 people, not less than 20 cars per 10,000 people in big cities, not less than 5 cars in small cities and medium values in middle cities [9]), the sharing ratio of taxi in public transport (between 10 and 20%), mileage utilization or the rate of free travel (threshold of 60% [10], 65% [11], 70% [12]), the average waiting time of passenger (10 min in flat peak period, 15 min in peak period of Zhejiang [10]; The expected value of Ningbo city is 6.72 min and the actual value is 12.49 min [10]; Dalian city expectations for 3.75 min [11]), the revenue per 100 km or the average daily revenue of single taxi (related to cost expenditure, business volume and other factors, without specific quantitative criteria), the index of road traffic congestion or the average waiting time of single taxi's carrying (no clear threshold standard), the average daily loading times of single taxi (no threshold) and other indicators. Therefore, this paper based on the taxi information management system operating data, and data mining-decision tree method is used to construct the cruise taxi capacity scale adjustment regression tree model by considering these indexes. Finally, we put forward the cruise taxi dynamic adjustment mechanism, the threshold value and the ordering relation of key indicators to provide theoretical basis and decision support for city cruise taxi capacity scale adjustment.

2 Analysis of the Influence Factors of Cruising Taxi Capacity

There are many influential factors in the cruise taxi capacity. The influential factors on the macro level includes the nature and functions of the city, tourist attraction, the development of socio-economic, the composition of population, the size and layout of city, natural geographical conditions, the infrastructure of urban transport and the quality of bus service. Micro-level, the taxi ownership per person, the sharing ratio of taxi in public transport, mileage utilization or the rate of free travel, the average waiting time of passenger and other indicators are the key indicators to describe

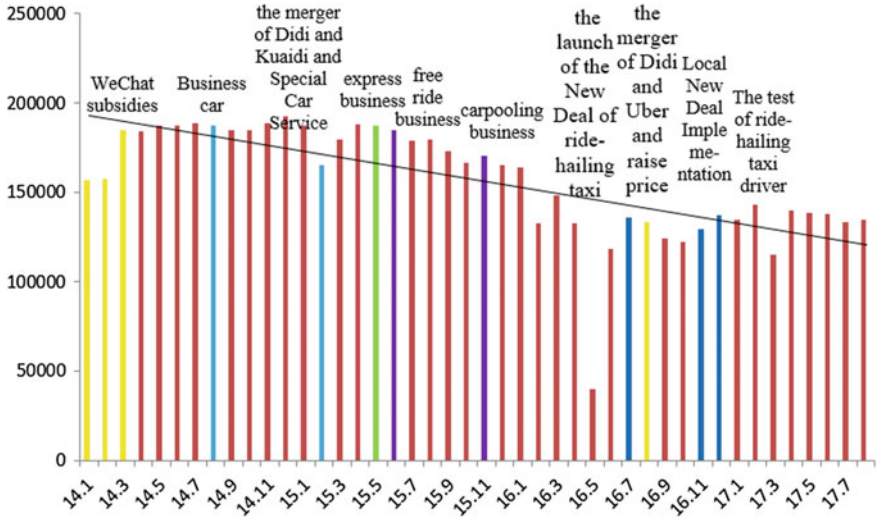


Fig. 1 Relationship between total monthly loading times of cruising taxi and marketing campaigns of ride-hailing taxi

matching degree between supply and demand of cruise taxi, which represent the capacity scale is enough or not. Based on the Ningbo City's cruising taxi operation data, this paper focuses on discussing the impact analysis of the ride-hailing taxi on the operation efficiency of the cruising taxi.

As shown in Fig. 1, the changes of the average monthly total vehicle number of the cruise taxi under Didi's different marketing activities and key businesses launches from 2014 to 2017 are described. From January to March 2014, Didi did not launch special car, express car and other business, just rely on the cruise car to seize the market. Among them, Didi has cooperated with WeChat, WeChat subsidy has increased its daily order volume from 350,000 to 5,218,300, with a market share of 60.2%. With the introduction of the special car business, the cruise taxi relied on the platforms such as Didi to increase the matching degree of supply and demand, and the operation volume has not been impacted, which still maintained a high level; However, with the merger of Didi and Kuaidi, and the launch of express, free ride and carpooling business, Didi and other ride-hailing taxi platforms already do not need cruise taxi to increase market share. A large number of private cars into the passenger transportation market by joining Didi and other ride-hailing taxi platforms (Shenzhou car and other platforms have their own vehicles), the barbaric growth of ride-hailing taxi and disruptive subsidies wars led to the business level of cruising taxi decreased significantly; Until the launch of the New Deal of ride-hailing taxi, the merger of Didi and Uber and the implementation of its price-raising plan, the downward trend in the business level of the cruise taxi was suppressed. With the speed-up of the ride-hailing taxi's legalization process, the overall business volume of the cruise taxi has generally stabilized.

3 Regression Tree Model and Algorithm

3.1 Regression Tree Model

There are many indicators that characterize whether the capacity of cruising taxi is sufficient or not. It is inappropriate to adjust the cruising taxi capacity scale by use a single standard, and the weight system evaluation method is difficult to give a clear ranking and coupling relationship among the indicators [13]. However, the decision tree, as a hierarchical data structure for implementing the divide-and-conquer strategy, adopting the prioritized strategy to continuously segment and output the forecast results from top to bottom, and can be converted into simple rules that are easy to understand [14]. Among them, the regression tree uses the impurity of regression as the standard of division. Assuming that a key indicator such as mileage utilization corresponds to decision node m and X_m is a subset of X to node m, that it is all x that satisfies all the decision point conditions from the root to node m of x ∈ X, then

$$b_m(x) = \begin{cases} 1, \ if \ x \in X_m : x \ arrival \ node \ m \\ 0, \ otherwise \end{cases} \tag{1}$$

which, $b_m(x)$ is decision node classification results, X_m is decision tree variables collection, m is decision tree node, x is a subset of the categories in the decision tree collection.

The mean square error of the estimated value is the partitioning metric, and g_m is the estimated value in the node m,

$$E_m = \frac{1}{N_m} \sum_t (r^t - g_m)^2 b_m(x^t) \tag{2}$$

Which $N_m = |\chi_m| = \sum_t b_m(x^t)$

Which, E_m is mean square error, N_m is sample size, g_m is the estimated value of node m, χ_m is number of training samples.

In the node, use the mean value of the output which required to reach the instance of this node:

$$g_m = \frac{\sum_t b_m(x^t) r^t}{\sum_t b_m(x^t)} \tag{3}$$

Which, r^t is input value, t is serial number, r is the input values, E_m corresponds to the variance of m. If the error is acceptable ($E_m < \theta_r$), then a tree node will be created to store g_m; Otherwise, the data of the node m can be further divided, making the sum of error of the branch minimum. On each node, founding the partition threshold of the attribute of minimizing error and numeric, and then the above process is performed recursively.

Let X_{mj} be a subset of the branching j for X_m:$\bigcup_{j=1}^{m} X_{mj} = X_m$, define

$$b_{mj}(x) = \begin{cases} 1, \ if \ x \in X_{mj} : x \ reaches \ the \ node \ m \ and \ takes \ the \ branch \ j \\ 0, \ otherwise \end{cases} \tag{4}$$

Let g_{mj} be the estimated value of the branch j that reaches the node m

$$g_{mj} = \frac{\sum_t b_{mj}(x^t) r^t}{\sum_t b_{mj}(x^t)} \tag{5}$$

And the error after division is

$$E_m^{'} = \frac{1}{N_m} \sum_j \sum_t (r^t - g_{mj})^2 b_{mj}(x^t) \tag{6}$$

which, $E_m^{'}$ is the mean square error after division, j is branch subset.

For arbitrary division, the reduction of the error is given by the difference between the Formula (2) and the Formula (6). To find the division direction of the reduction of the maximization error or equivalent to take the minimum of the Formula (6). The mean square error is a possible error function, Another maximum error

$$E_m = \max_j \max_t |r^t - g_{mj}| b_{mj}(x^t) \tag{7}$$

The maximum error can be used to ensure that the error of any instance is not greater than the given threshold. The acceptable error threshold is a function of complexity; The smaller the value, the larger the tree and the greater the risk of overfitting, the greater the value, the greater the likelihood of insufficient fitting and excessive smoothness.

When each decision node uses all key indicators as input dimension, the linear multivariable node is defined as

$$f_m(x) : w_m^T x + w_{m0} > 0 \tag{8}$$

which, $f_m(x)$ is multivariate tree functions, w_m^T is weight transposed matrix, w_{m0} is threshold matrix.

Instances are further divided from successive nodes on the path from root to leaf, while leaf nodes define polyhedrons on the input space, fine-tune the weights w_{mj} one by one to reduce the statistical significance indicators, through subset selection to reduce the dimension and the complexity of the node.

3.2 Model Algorithm

According to the characteristics of the taxi management information system data structure, the data types of variables and explanatory variables are numeric, and

the Chi-squared Automatic Interaction Detector (CHAID) of variance analysis is used as the algorithm to construct regression tree. Use the input variables of initial grouping as the control variables of the variance analysis, the output variables as observational variables, then examining whether there is a significant difference in the distribution of output variables under the initial grouping of input variables, carrying out multiple comparison tests and merging adjacent groups which have not significant distribution difference. Repeat until there are no merged groups. For the preprocessed input variables, calculating the probability-p value of the statistic associated with the output variable correlation test. For numeric output variables, use F statistic. The input variable with the lowest probability-p value is most closely related to the output variable and should be the current best group variable. When the probability-p values are the same, we should choose to test the input variables with the largest statistics observed value, no longer redefine the segmentation points, and every category of the grouping variables is automatically classified as a tree branch to form a multi branch tree.

The model risk assessment index of the regression tree is mainly the variance and misjudgment rate of the predicted value. The standard error $\propto_w$ of the risk trajectory is

$$\alpha_w = \sqrt{\frac{p(1-p)}{N}} \tag{9}$$

which, p is risk estimates, N is Sample size.

We use maintenance method to evaluate the performance of regression tree model and divide the labeled raw data into two disjoint sets, which are training set and test set respectively. Summarize the model on the training dataset and evaluate the performance of the model on the test set.

4 Structure and Result Analysis of Decision Tree

Regression tree method was used to classify the operational data and use the capacity scale prediction value as the target attribute, to supervise the classification of the samples. The confidence level of the model is 87.36%, which indicates that the cruise taxi capacity scale adjustment threshold value is well summarized and explained. As shown in Fig. 2, the regression tree is a multi-tree structure with a depth of 3, the root node is the forecast value of capacity scale, rectangles represent as leaf nodes. The decision variables, the number of samples, the mean of sample and forecast value, F statistic value and probability-p value are given for each node respectively. The influence factors such as mileage utilization, the average daily revenue of single taxi, revenue per 100 km, are brought into the model in order by relevance descending, indicating that the scale of capacity was significantly affected. Among them, mileage utilization has the strongest effect on the scale of cruising taxi capacity and F statistics was 63.73; Followed by the average daily revenue of single taxi, the average waiting

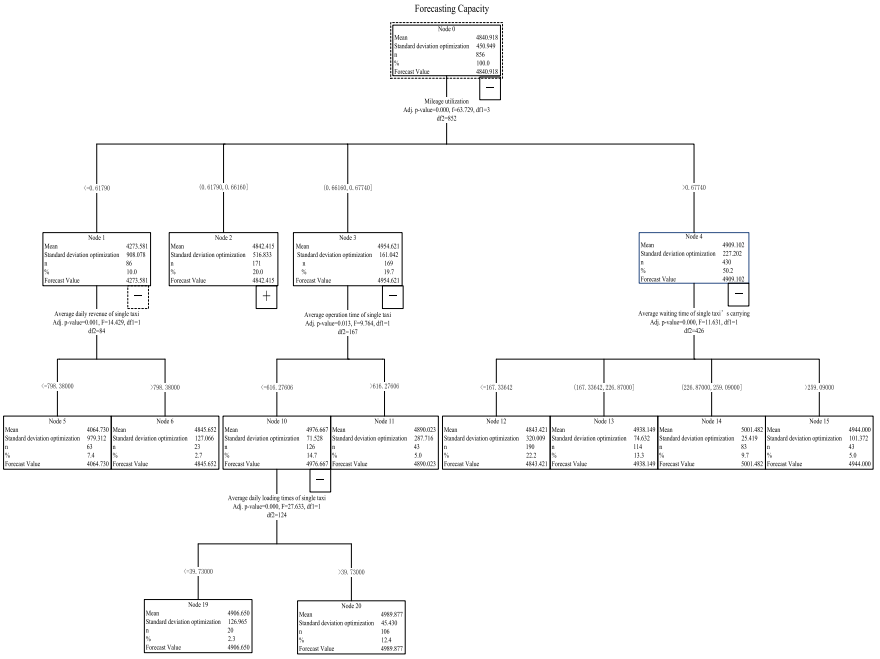


Fig. 2 Regression tree structure of the scale of the cruising taxi

time of single taxi's carrying, the average operation time of single taxi, and revenue per 100 km. The statistics of F decreases in order, indicating that the degree of influence degressive; Finally, the average daily loading times of single taxi is chosen as the third layer of classification indicator; *p*-values of the indicators in every layer are less than 0.05 and all pass the significance tests; It can be seen that the results of tree structure analysis reveal the interaction, ranking and coupling relationship among various influencing factors, and can specifically analyze the role of each threshold in sub-groups, and provide a richer and more precise threshold criterion for the capacity scale adjustment and delivery.

According to the result of decision tree analysis, 14 rules are generated. Combined with the interactive relationship between the price and capacity of the cruise taxi, the following summarizes the characteristics of the scale adjustment and delivery of the cruise taxi capacity as follows:

Node 5, the mileage utilization is low, indicating that the cruise taxi oversupply and its operating efficiency is low; The average daily revenue of single taxi is low, indicating that the driver's income is low; Under this threshold value condition, the actual capacity is oversize, and need to reduce capacity ($4427-4065 = 362$ vehicles, 4427 is the number of cruising taxi in Ningbo since 2013) so as to achieve a balance between supply and demand.

Node 6, the mileage utilization is low, but the average daily revenue of single taxi indicates that driver's income is acceptable. Under this threshold value condition,

capacity scale does not to be adjusted, but in order to improve mileage utilization, the passenger flow attraction rate can be increased through price leverage (price reduction).

Node 7, the mileage utilization is within the range of (0.6179, 0.6616], and the revenue per hundred kilometers is less than or equal to 225.01, slightly less than the actual basic freight rate value (2.35 yuan/km), which shows that the operate condition of cruise taxi is poor. However, the capacity prediction value of this decision node is slightly lower than the parent node 2, capacity scale adjustment window does not open, should adapt to the basic freight rates value under the condition of this threshold value condition and aim at improving revenue level, and to adapt to the market situation in this stage of supply and demand by reducing freight.

Node 16, mileage utilization and the revenue per hundred kilometers are all in an acceptable range, but the average daily loading times of single taxi is low, which indicates that passenger transportation demand is insufficient. Therefore, the capacity adjustment window does not open, and the passenger demand can be improved through price cuts.

Node 17, all indicators are in an acceptable range, capacity scale adjustment window does not open, no need to adjust the price.

Node 18, contrary to node 16, the average daily loading times of single taxi is high, indicating that the passenger demand is strong. Therefore, the capacity adjustment window does not open, and the passenger demand can be suppressed by the price increase.

Node 9, contrary to node 7, the revenue per 100 km is more than 235.8, slightly higher than the actual basic freight rate value (2.35 yuan/km), the capacity adjustment window does not open, and it can adapt to the high-price market conditions by increasing the price.

Node 19, the mileage utilization is within the range of (0.6179,0.6616], the threshold value is high, indicating that demand is slightly larger than the supply, but the average operation time of single taxi is less than or equal to 616.28, indicating that the driver's working strength does not exceed the limit of tolerance; And the average daily loading times of single taxi is less than or equal to 39.73, does not exceed the limit value, indicating that the difference between demand and supply is acceptable, there is no need to adjust the scale of capacity to achieve a balance between supply and demand.

Node 20, contrary to node 19, the average daily loading times of single taxi exceeds the limit value, indicating that there is a large difference between demand and supply. In order to achieve the balance between supply and demand, drivers may be appropriately encouraged to extend their working hours to reduce mileage utilization, the average daily loading times of single taxi and passenger demand suppression by increasing the price appropriately.

Node 11, similar to nodes 19 and 20, mileage utilization is high, but the average operation time of single taxi exceeds the limit value. The driver's working intensity is too high, which is detrimental to healthy development of the taxi industry. Therefore, the adjustment window is opened and delivery capacity ($4890 - 4427 = 463$ vehicles) to achieve the goal of reducing mileage utilization and driver's working intensity.

Node 12, mileage utilization is greater than 0.6774, the threshold value is in a high level, indicating demand is greater than the supply. But the average waiting time of single taxi's carrying is less than or equal to 167.34, indicating that the impact of urban traffic congestion is smaller, the service quality of cruise taxi under the condition of high mileage utilization is still in a high level, it is not recommended to adjust capacity scale.

Node 13, the average waiting time of single taxi's loading increases, and traffic congestion has a greater impact, which increases the waiting time of cruise taxis service, and reduces the service quality. In order to make up for low speed taxi costs, it is suggested that a waiting fee be added to the freight rate standard.

Node 14, the average waiting time of single taxi's loading is further increased, and the traffic congestion deteriorates, taxi needs to pay extra cost for completing service under the condition of high mileage utilization. Therefore, it is recommended to raise waiting rate to make up for extra cost.

Node 15, the average waiting time of single taxi's loading exceeds the patience value, it shows that the traffic condition of the city has deteriorated drastically and taxi needs to pay extra cost for completing service under the condition of high mileage utilization. At this time, capacity scale adjustment window is opened, put in appropriate capacity ($4812 - 4427 = 385$ vehicles) to reduce the mileage utilization, make up for low-speed driving costs and reach a new balance of supply and demand.

In addition, as shown in Fig. 3, the accumulative returns of training samples and test samples are given respectively, which has achieved a high cumulative income in the early stage, and then quickly approached 100% and smoothed, indicating that the regression tree model is ideal. The variance and misjudgment rate of prediction value were used to evaluate the risk of regression tree model. The standard error of training and test samples was 6.13%, and the misjudgment rate was 0.07. The overall accuracy of the model was better.

5 Conclusion

Due to the lack of scale adjustment index system of cruising taxi capacity, this paper is based on the taxi operation datasets from Taxi Information Management System, and Regression Tree model for the scale adjustment of cruising taxi capacity was established by considering the average daily loading times of single taxi, mileage utilization, the level of revenue and other indicators, obtained the corresponding threshold value interval. Then the dynamic adjustment mechanism of the cruising taxi capacity scale and the thresholds of the key indicators were proposed by coupling the representation functions on the balance between supply and demand with the importance ordering relationship of indicators, which consisted of the taxi ownership per person, the sharing ratio of taxi in public transport trip structure, the sharing ratio between cruising and ride-hailing taxi, and the average waiting time of passengers. This dynamic adjustment mechanism of the cruising taxicab capacity scale and the thresholds of the key indicators provide the theoretical basis and decision support

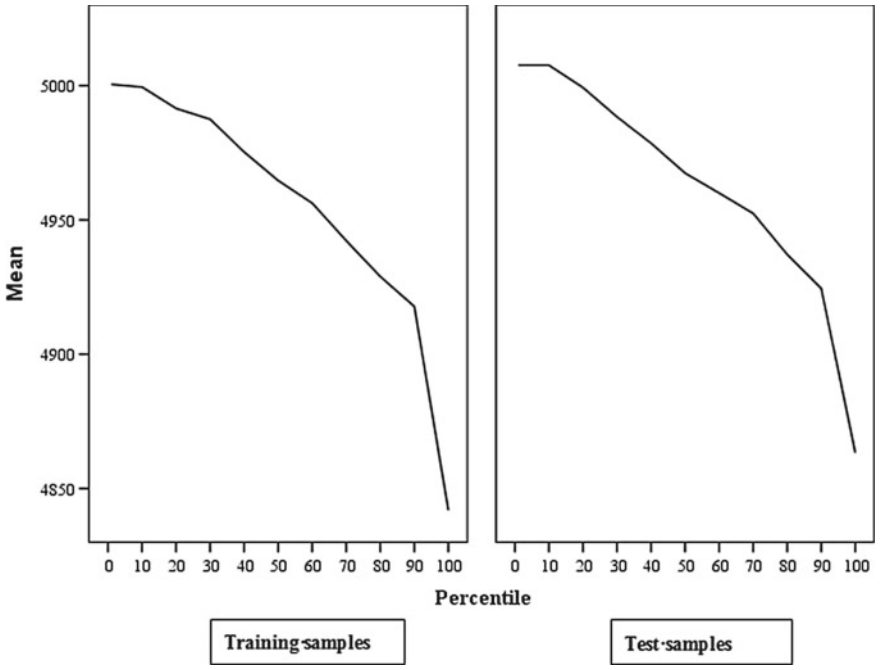


Fig. 3 Cumulative income figures of training and testing samples

for the scale adjustment of the urban cruising taxi capacity. The results indicate that, (1) in the three-layer structure of regression tree, mileage utilization has the strongest effect on the cruising taxicab capacity scale, with a F statistic of 63.73; then the average daily revenue of single taxi, the average waiting time of single taxi's carrying, the average operation time of single taxi, and the revenue per 100 km have the stronger influence which decreased progressively; And the average daily loading time of single taxi is chosen as the third layer of classification indicator, *p*-values of the indicators in every layer are less than 0.05 and all pass the significance tests; (2) When mileage utilization ratio is less than 0.6179 and the average daily revenue of single taxi is less than 798.38 Yuan, the scale of cruising taxi capacity is surplus and need to reduce 362 vehicles; (3) When mileage utilization is between 0.6616 and 0.6774 and the average operation time of single taxi is more than 616.28 min, also mileage utilization is more than 0.6774 and average waiting time of single taxi's carrying is more than 259.09 s, the scale of cruising taxi capacity is insufficient and 385 vehicles and 463 vehicles are proposed to put in respectively; (4) The threshold value of the taxi ownership per person, the sharing ratio of taxi in public transport trip structure, the sharing ratio between cruising and ride-hailing taxi, and the average waiting time of passengers are 23 vehicles per ten thousand people, 18%, 10:6.54 and 11.38 min respectively.

Acknowledgements This research is supported by National Natural Science Foundation of China (No. 51408322 and 71701108) and Natural Science Foundation of Zhejiang Province (LY20E080011).

References

1. Ch JDC, Briones J (2006) A diagrammatic analysis of the market for cruising taxis. Transp Res Part E 42(6):498–526
2. Sheng C (2015) Impact of "Internet + Travel" on the Traditional Taxi Industry. Foreign Econ Relat Trade 255(9):76–77
3. Xiao Y, Lang WEI (2016) Research on the reform path of taxi price. Price: Theory Pract 23(7):79–81
4. Office of the State Council (2016) Guiding opinions on deepening reform to promote the healthy development of the taxi industry. Urban Plan Commun 815(15):13
5. Luan Lina (2015) Taxi operation analysis and data mining based on urban taxi GPS Data. Shandong University, Jinan
6. Douglas GW (1972) Price regulation and optimal service standards: the taxicab industry. J Transp Econ Policy 6(8):116–127
7. Yang H, Ye M, Tang WH (2005) Regulating taxi services in the presence of congestion externality. Transp Res Part A 39(1):17–40
8. Skok W, Martinez JA (2010) An International taxi cab evaluation. Knowl Process Manag 34(3):112–127
9. Ministry of Construction of the People's Republic of China. GB50220-95 (1995) Urban road traffic planning and design specification. China Planning Press, Beijing
10. Yang Ren-fa (2006) Study on building of supervision system of urban passenger transport taxi. Transp Stand 32(6):122–126
11. Liu Hong-ting (2011) Assessment and optimization of the scale of taxi capacity. Dalian Maritime University, Dalian
12. Cao Y, Luo X (2015) Taxi service network equilibrium model with taxi-hailing apps. J Chang'an Univ (Nat Sci Ed) 35(1):203–208
13. Wong KI, Wong SC, Yang H (2001) Modeling urban taxi services in congested road networks with elastic demand. Transp Res Part B 35(9):819–842
14. Xue W, C Huan-ge (2010) Method and application of clementine data mining. Publishing House of Electronics Industry, Beijing, pp 149–162

Impact of Road Alignment on Lane Departure: A Driving Simulator Study

Weiwei Guo, Mengqi Ren, Jiyuan Tan and Yan Mao

Abstract Lane departure is a major cause of side-swipes, rear-end collisions, and crashes. Road curvature, as an important part of linear road design, plays a vital role in lane departure. Based on study of the Curve radius and slope of the road, the linear parameters can be optimized to reduce the probability of lane departure. The curvature radius and slope of the road are designed in UC-winRoad and multiple experiments through a driving simulator are performed to output the evaluation data. Based on the radius of curve, slope, and turning direction, numerical analysis of the lane offset are proposed. The relationship between lane departure and road alignment is obtained by analyzing the |P1| value after fitting the vehicle trajectory. The study found that the right-turn curves have a larger lane departure than the left-turn curves, and that too small and too large a curve radius will result in a large lane departure.

Keywords Driving simulation · Road alignment · Road curvature · Traffic safety

1 Introduction

Traffic accidents have become a major issue that threatens the safety of the public. Poor road conditions are a potential cause of traffic accidents. The pros and cons of road conditions mainly refer to the liner of the road, the radius of the curve, the slope and width of the road, the roadbed and the road surface, etc. According to the different external objective conditions, road design does not have a reasonable and safe unified standard. On the other hand, road design should not only consider external conditions but also consider the influence of the combined alignments, the visual impact of the road itself on the driver, and so on. These problems are of great significance in achieving better road conditions.

W. Guo · M. Ren · J. Tan (✉)
Beijing Key Lab of Urban Intelligent Traffic Control Technology,
North China University of Technology, Beijing 100144, China
e-mail: tjyphilip@163.com

Y. Mao
Research Institute of Highway Ministry of Transport, Beijing 100088, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_61

Studies by Eustace et al. [1], Lord et al. [2], and Liu and Subramanian [3] have shown that certain road alignments increase the likelihood of lane departure crashes. Therefore, poor road alignments may induce lane departures and cause traffic accidents. If the combination of horizontal curvature and the slope is not properly designed, it may affect the lane keeping of a vehicle.

In addition, since the driver is located on one side of the vehicle, the distance between the right and left rear view mirrors and the left and right side lanes of the current lane is different, and the influence of the different turning directions on the driver is also inconsistent. Chen et al. [4] showed that the drivers on left-turn curves are biased to the left, and the drivers on right-turn curves are biased to the right. Such indications point to a need for direction of turning to also be considered in the linear design of the road. The design criteria for the combined alignments should also be adjusted according to the direction of turning.

The Design Specification for Highway Alignment [5] focuses on determining the thresholds for individual horizontal routes and individual vertical routes. Several qualitative design guidelines for combined alignments are also proposed. However, the current guidelines do not systematically consider the safety standards of combined alignments. Moreover there are currently no different standards for different turning directions.

Therefore, the purpose of this paper is to explore the rule of curve radius, slope and turning direction as they affect lane departure in combined alignments, with an aim toward guiding linear design. Since the corresponding lane departure data is less likely to be obtained under real-world studies, a driving simulator is used for research. The distance of the vehicle from the centerline of the lane, the position of the vehicle in the coordinate system and the distance of the vehicle from the starting point are continuously captured by the simulator. Factors such as the road environment and traffic conditions are consistent in the simulation to reduce the external impact on lane departures. Statistical methods and linear regression analyses were used to analyze the distance of the vehicle from the centerline of the lane and the fitting vehicle trajectory.

2 Literature Review

This section will summarize the existing related research, including the impact of road alignment on lane departures and methods for assessing these effects.

2.1 Road Alignment Effects on Safety

Horizontal curvature and slope grades have been found to be associated with accidents in many studies [6–10]. A poor design of the combined alignments will have a negative impact on the driver and result in traffic accidents [11, 12]. On the one

hand, it will increase the driver's load and increase the difficulty of driving; on the other hand, it will lead to the driver's false perception of the geometric design of the curved part, and inappropriate driver behavior that differs from the design expectation will arise [13].

Lamm et al. [14] demonstrated that combined alignments can affect the safety of highways. In the 1970s, there were some reports in the literature on the relationship between driver trajectory and road geometry design [15, 16]. By identifying the relevant variables of a crash, Abdel-Aty et al. [17] found that the crash rate was strongly influenced by road curvature and uphill or downhill segments. Hanno [18] studied the relationship between vertical and horizontal curves and found that the combined alignments affected the crash rate. However they did not consider the direction of the curve.

2.2 Road Alignment Effects on Lane Departure

Crash data is difficult to obtain in a particular experiment, so surrogate measures are used that are directly related to traffic accidents and are affected by road alignment [19, 20].

Lane departure and turning trajectory are also excellent indicators for evaluating the linear design of a road. Wu et al. [21] used the multivariate linear regression model to construct the standard deviation prediction model of lateral offset, and the model showed that the tangent length, the length of the circular curve and the rate of curvature change are important independent variables that affect the lateral offset. Lin et al. [22] found that the greater the curvature, the greater the lateral offset. Said et al. [23] studied the driver's performance on the curved road through road linearity and vehicle position, and found that the vehicle trajectory was inconsistent with the actual curve path. The study by Spacek [24] shows that the distance between the vehicle position on the curve and the roadside is significantly greater than a straight line, and the vehicle position on the left-turn curve is closer to the lane centerline than on the right-turn curve. Research on the impact of road design on lane departures has focused on horizontal parameters without considering the effects of vertical design. Chen et al. [4] studied how the combined alignments affected different offset events by modeling a mixed multi-logarithmic model, but this study did not consider the slope.

Based on the results and shortcomings of previous studies, this paper comprehensively considers the factors of curve radius, slope and curve direction on road alignment, and analyzes the lane departure and vehicle trajectory to obtain more accurate conclusions. The North China University of Technology driving simulator was used in this study, and the experimental data was collected by a computer in real time [25].

3 Methodology

A four-lane (two lanes in each direction) expressway was simulated using the North China University of Technology driving simulator, which consisted of 72 vertical and horizontal combined alignments. All participants experienced each of these alignments while their lane departure and position data were continuously acquired.

3.1 Participants

Five drivers (3 males and two females) who had each obtained a driver's license and had different types of driving experience served as the participants. Their average age was 23 years old. Because there is a certain gap between the driving simulator and a real driving experience, the participants were trained before the formal experiment, so that the participants could be familiar with the operation of the driving simulator and could be as smooth and skilled as in actual driving.

3.2 Simulated Road Course Design

This paper mainly studies the lane departure under different combinations of curvature radius and slope. In the real world, the radius of the circular curve is 400–1500 m, and a curve is taken every 100 m. The longitudinal slopes are 1, 3, and 5%, which include two curve directions (left-turn and right-turn).

The curve of the radius of the same circular curve is set as the adjacent two reverse circular curves. Due to the requirements for S-shaped curves in China's highway route design specification, a sufficiently straight line length is designed between the two reverse circular curves [1] (see Fig. 1). The total length of the road is 63.39 km, with four–lanes (two lanes in each direction), and the designed speed is 100 km/h (see Fig. 2).

Due to the long length of the road required for the experiment, in order to reduce the length of the road design and to take into account the actual situation, the slope is only designed to be 1, 3 and −5%. All the data was obtained from the starting point to the end and from the end to the starting point in the simulated driving (see Fig. 3).

3.3 Experimental Procedures

In this experiment, participants drive on the established road model in the driving simulation environment, and the lane departure when cornering is evaluated by the real-time data output for the vehicle's motion state. Participants have 5 to 10 min to

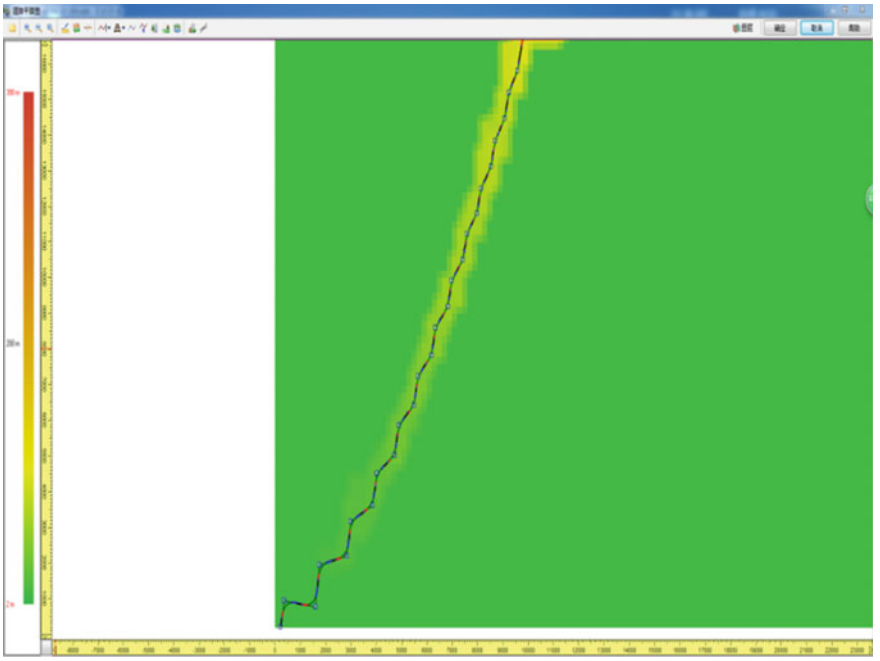


Fig. 1 Road plane

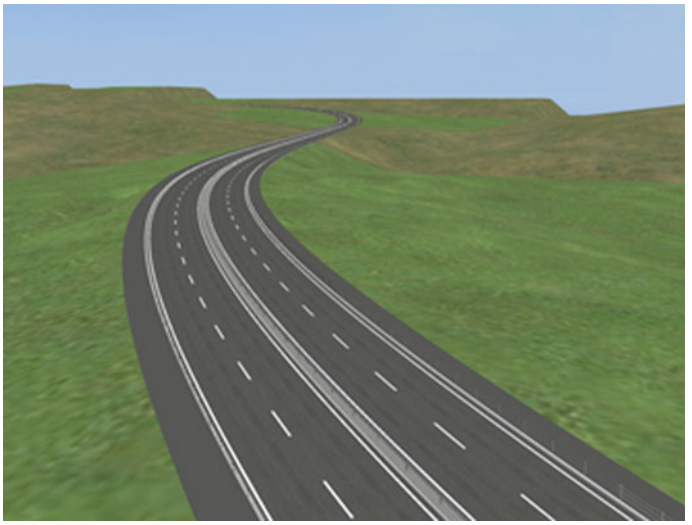


Fig. 2 Road model

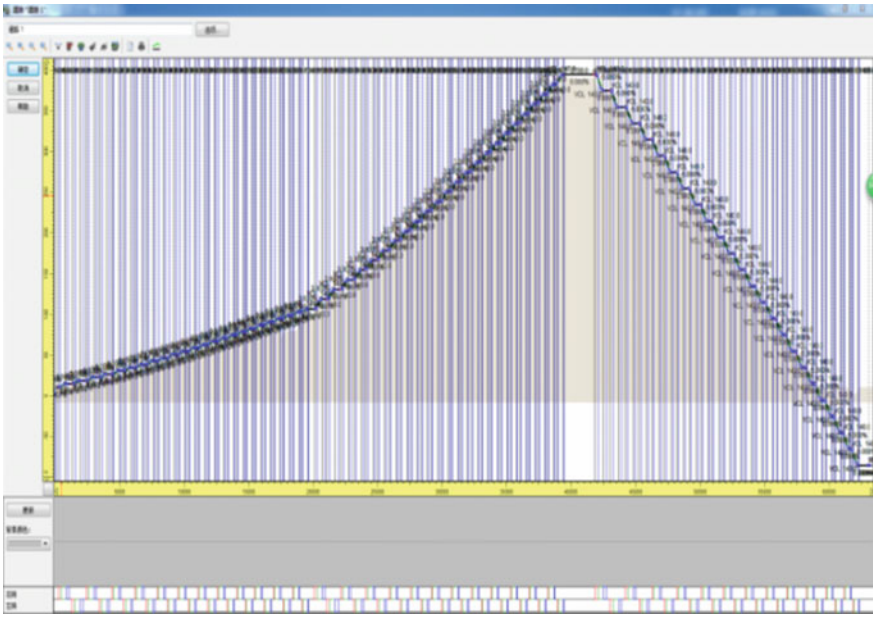


Fig. 3 Road longitudinal section

Fig. 4 Driving simulation experiment

adapt to the driving simulator before the experiment. The speed is always maintained at 100 km/h during driving, and can be adjusted slightly according to the actual situation during a turn (see Fig. 4).

4 Results and Analyses

Based on purpose of the research, the distance of the vehicle is determined by Distance Along Road (the distance from the starting point of the road), and the vehicle offset is numerically analyzed by Offset from Lane Center (the distance of the vehicle position from the center line, hereinafter referred to as the lane departure). Position X and position Z (vehicle position) establish a driving trajectory model and analyze its trajectory shape. We believe that the combination of these two analyses can more accurately reveal the impact of the combined alignments on lane departure.

4.1 Lane Departure

4.1.1 Effect of Radius of the Curve on Lane Departure

The initial experimental data is obtained by averaging the lane departure at all time points in a single curve based on the raw data. All the experimental data obtained were taken as absolute values and divided into six groups according to the gradient. Table 1 the lane departure for the same circular curve radius and the directions are averaged, and then the average value is summed according to the left-turns and the right-turns.

By comparing the influence of radius of curvature on the lane departures during the turning process, it is easy to see that the lane departure decreases first and then increases with the increase of the radius of the circular curve, indicating that the lane departure increases with the increase of the radius of the circular curve. Yet when the curve radius exceeds 1000 m, it is positively correlated with the lane departure (see

Table 1 Lane departure for different radii

Curve radius/m	Right-turn/m(SD)	Left-turn/m(SD)	Sum/m
400	0.288(0.211)	0.245(0.150)	0.533
500	0.254(0.169)	0.239(0.149)	0.494
600	0.246(0.193)	0.199(0.171)	0.445
700	0.264(0.224)	0.192(0.134)	0.456
800	0.228(0.182)	0.198(0.146)	0.425
900	0.223(0.164)	0.196(0.133)	0.419
1000	0.161(0.145)	0.212(0.163)	0.373
1100	0.180(0.145)	0.225(0.188)	0.405
1200	0.189(0.129)	0.187(0.186)	0.376
1300	0.229(0.174)	0.281(0.207)	0.511
1400	0.230(0.188)	0.219(0.163)	0.449
1500	0.257(0.165)	0.249(0.167)	0.506

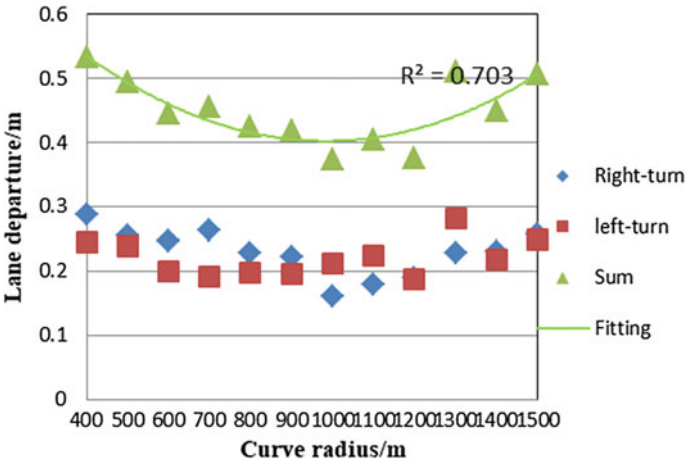


Fig. 5 Lane departure at different radii

Fig. 5). Fit the sum of the left-turn and right-turn data (where the brackets include the 95% confidence intervals for all parameters):

$$y = 3.923 \times 10^{-7} R^2 - 0.0007716R + 0.7816 \quad (1)$$

$$(1.954 \times 10^{-7}, 5.891 \times 10^{-7})(-0.00115, -0.0003928)(0.6159, 0.9474)$$

4.1.2 Effect of Slope on Lane Departure

In order to analyze the influence of the slope on the lane departure, all data with the same turning direction and slope grade are averaged, and then the average value is summed according to the left-turn and the right-turn in Table 2.

According to the lane departures under different slopes, it can be clearly found that whether it is uphill or downhill, as the slope becomes larger and larger, the lane

Table 2 Lane departure for different slope

Slope (%)	Right-turn/m(SD)	Left-turn/m(SD)	Sum/m
1	0.265(0.177)	0.223(0.179)	0.488
−1	0.261(0.155)	0.218(0.140)	0.479
3	0.240(0.162)	0.273(0.182)	0.513
−3	0.233(0.158)	0.293(0.153)	0.526
5	0.337(0.232)	0.234(0.147)	0.571
−5	0.314(0.165)	0.307(0.183)	0.621

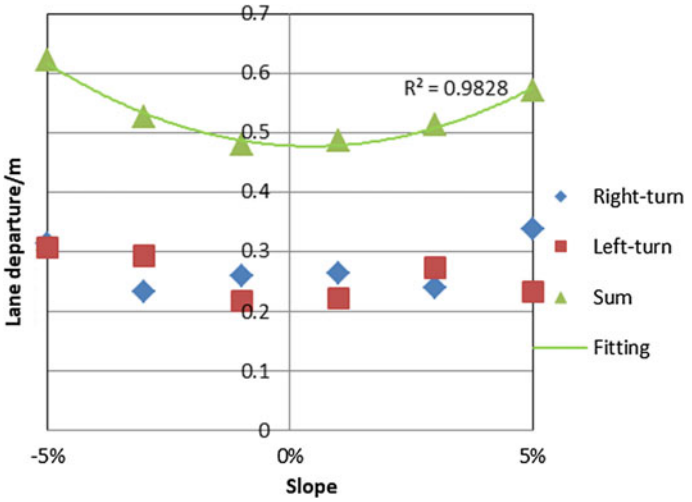


Fig. 6 Lane departure at different slopes

departure during the turning process become increasingly larger (see Fig. 6). Fit the sum of the left- and right-turn data (where the brackets include the 95% confidence intervals for all parameters):

$$y = 47.18G^2 - 0.4095G + 0.478 \tag{2}$$

$$(35.22, 59.15)(-0.759, -0.05996)(0.4596, 0.4963)$$

4.1.3 Effect of Curve Direction on Lane Departure

From the previous analysis, it can be seen that the slope and curve radius have different effects on the lane departure in different curve directions. Note that the direction of the curve affects the lane departure.

4.2 Turning Track

In order to more intuitively observe the lateral change of the vehicle, the trajectory of the vehicle during the turning is drawn according to the vehicle's position in the UC, and is fitted as a quadratic function by MATLAB (see Fig. 7) (The mean of the mean square error of all fitted curves is 32.56):

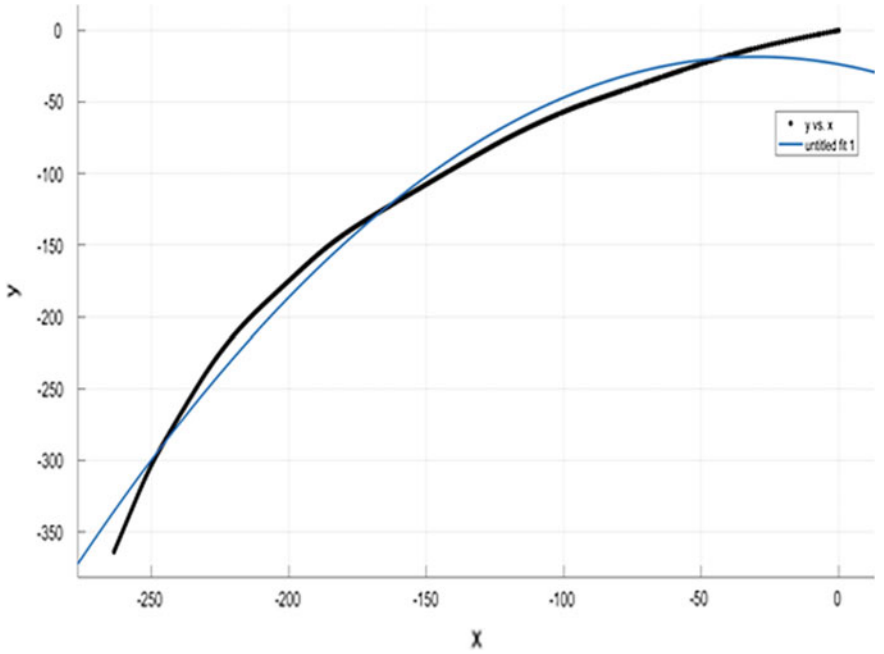


Fig. 7 Vehicle left-turn track (curve radius 1500 m, slope 5%)

$$y = P_1x^2 + P_2x + P_3 \tag{3}$$

P1 determines the shape of the quadratic curve, and the degree of bending of the curve can be reflected. It can be considered that by analyzing the change in the value of P_1, it is possible to investigate the variation in the trajectory of the vehicle when driving through different curves. The larger the $|P_1|$, the narrower the curve fitting, indicating that the vehicle is moving more gradually and closer to a straight line.

In order to explore the regular pattern of vehicle trajectory variation with the curve radius, $|P_1|$ is fitted with the curve radius as the independent variable, and is fitted to the quadratic function according to the data trend.

The diversification for the $|P_1|$ value of the right-turn at different slopes during the uphill process with the radius of the curve (see Fig. 8) is found such that the fitting curve of $|P_1|$ is similar on the 1 and 3% curves. This means that the turning track is basically unaffected by the slope change when the slope is small. Thus, $|P_1|$ in the case of 1 and 3% is fitted to one curve (see Fig. 9).

The $|P_1|$ Fitting curve when the slope is 1% or 3% (right-turn) is:

$$|P_1| = -1.406 \times 10^{-8}r^2 + 3.077 \times 10^{-5}r - 0.0043 \tag{4}$$

$$(-1.892 \times 10^{-8}, -9.196 \times 10^{-9})(2.141 \times 10^{-5}, 4.012 \times 10^{-5})(-0.0084, -0.0002)$$

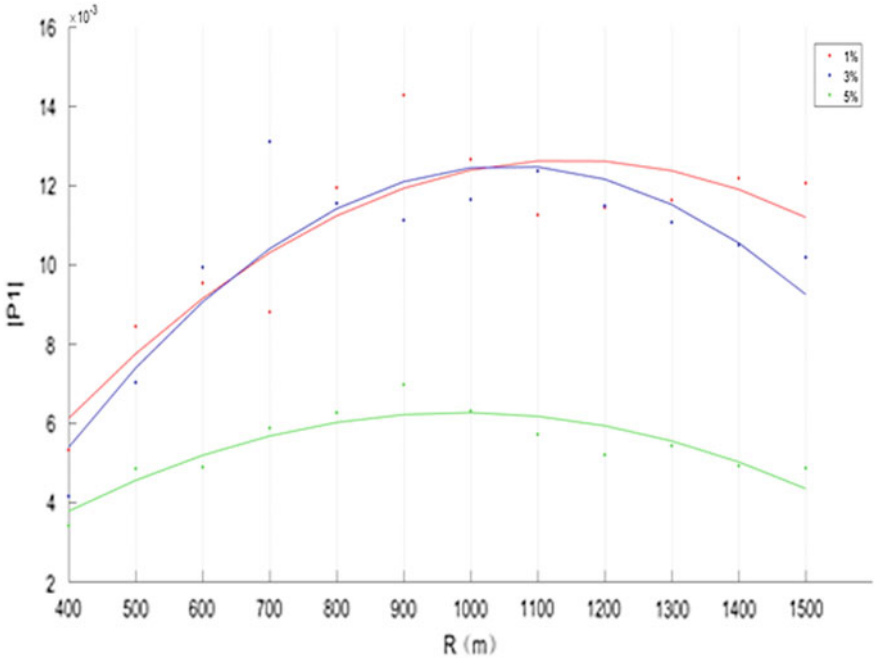


Fig. 8 Uphill and right-turn

where the brackets include the 95% confidence intervals for all parameters.

The $|P_1|$ Fitting curve when the slope is 5% (right-turn) is:

$$|P_1| = -7.247 \times 10^{-9} r^2 + 1.429 \times 10^{-5} r - 0.0008 \tag{5}$$

$$(-1.018 \times 10^{-8}, -4.311 \times 10^{-9})(8.635 \times 10^{-6}, 1.994 \times 10^{-5})(-0.0032, 0.0017)$$

where the brackets include the 95% confidence intervals for all parameters.

The diversification for the $|P_1|$ value of the left-turn at different slopes during the uphill process with the radius of the curve (see Fig. 10) is found such that the two fitting curves of $|P_1|$ are basically coincident on 1 and 3%. This means that the turning track is basically unaffected by the slope change when the slope is small. Thus, $|P_1|$ in the case of 1 and 3% is fitted to one curve (see Fig. 11).

The $|P_1|$ Fitting curve when the slope is 1% or 3% (left-turn) is:

$$P_1 = -1.137 \times 10^{-9} r^2 + 4.679 \times 10^{-6} r - 0.0004 \tag{6}$$

$$(-1.637 \times 10^{-9}, -6.38 \times 10^{-10})(3.718 \times 10^{-6}, 5.64 \times 10^{-6})(-0.0008, 0)$$

where the brackets include the 95% confidence intervals for all parameters.

The $|P_1|$ Fitting curve when the slope is 5% (right-turn) is:

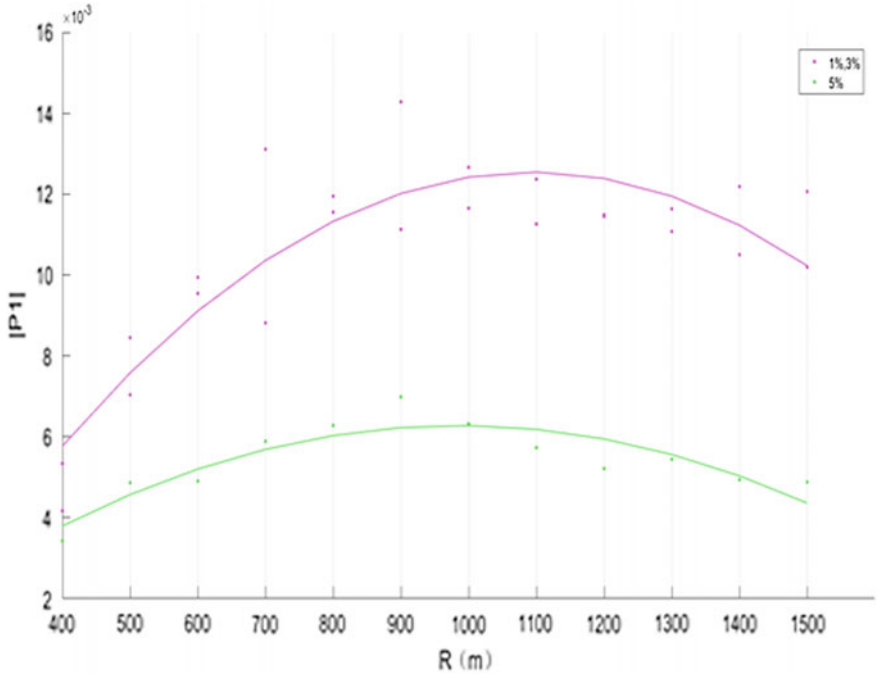


Fig. 9 Uphill and right-turn (fitting)

$$P_1 = -9.102 \times 10^{-10} r^2 + 2.706 \times 10^{-6} r + 0.0004 \tag{7}$$

$$(-1.697 \times 10^{-9}, -1.235 \times 10^{-10})(1.192 \times 10^{-6}, 4.22 \times 10^{-6})(-0.0003, 0.0011)$$

where the brackets include the 95% confidence intervals for all parameters.

The change in the P_1 value is similar to that of uphill or downhill (see Figs. 12 and 13).

The |P1| Fitting curve when the slope is -1% or -3% (right-turn) is:

$$|P_1| = -1.179 \times 10^{-8} r^2 + 2.557 \times 10^{-5} r - 0.002 \tag{8}$$

$$(-1.643 \times 10^{-8}, -7.151 \times 10^{-9})(1.664 \times 10^{-5}, 3.449 \times 10^{-5})(-0.0059, 0.0019)$$

where the brackets include the 95% confidence intervals for all parameters.

The $|P_1|$ Fitting curve when the slope is -5% (right-turn) is:

$$|P_1| = -8.253 \times 10^{-9} r^2 + 1.584 \times 10^{-5} r - 0.0009 \tag{9}$$

$$(-1.115 \times 10^{-8}, -5.36 \times 10^{-9})(1.027 \times 10^{-5}, 2.14 \times 10^{-5})(-0.0034, 0.0015)$$

where the brackets include the 95% confidence intervals for all parameters.

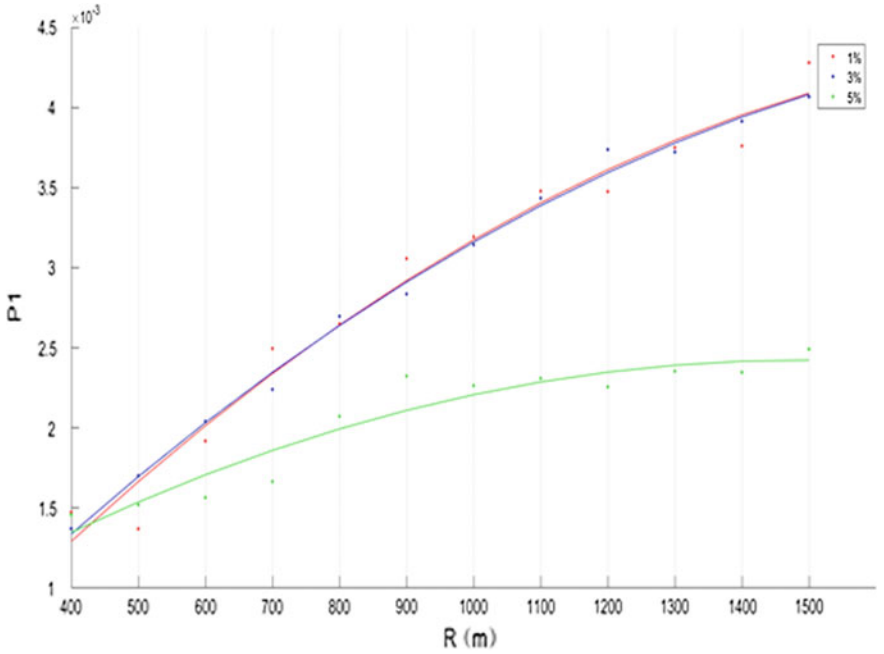


Fig. 10 Uphill and left-turn

The $|P_1|$ Fitting curve when the slope is -1% or -3% (left-turn) is:

$$P_1 = -9.309 \times 10^{-10} r^2 + 4.415 \times 10^{-6} r - 0.0002 \tag{10}$$

$$(-1.546 \times 10^{-9}, -3.16 \times 10^{-10})(3.232 \times 10^{-6}, 5.599 \times 10^{-6})(-0.0007, 0.0004)$$

where the brackets include the 95% confidence intervals for all parameters.

The $|P_1|$ Fitting curve when the slope is -5% (left-turn) is:

$$P_1 = -8.91 \times 10^{-10} r^2 + 2.508 \times 10^{-6} r + 0.0005 \tag{11}$$

$$(-1.67 \times 10^{-9}, -1.124 \times 10^{-10})(1.01 \times 10^{-6}, 4.006 \times 10^{-6})(-0.0002, 0.0011)$$

where the brackets include the 95% confidence intervals for all parameters.

4.2.1 Effect of Radius of the Curve on the Track

As the radius of the curve increases, $|P_1|$ gradually increases, but when the radius is greater than 1000 m, $|P_1|$ has a decreasing tendency. It shows that with the increase of the curve radius on the right-turn curve, the driving track tends to be gentle at first, and then the rear corner will gradually increase (see Figs. 9 and 12).

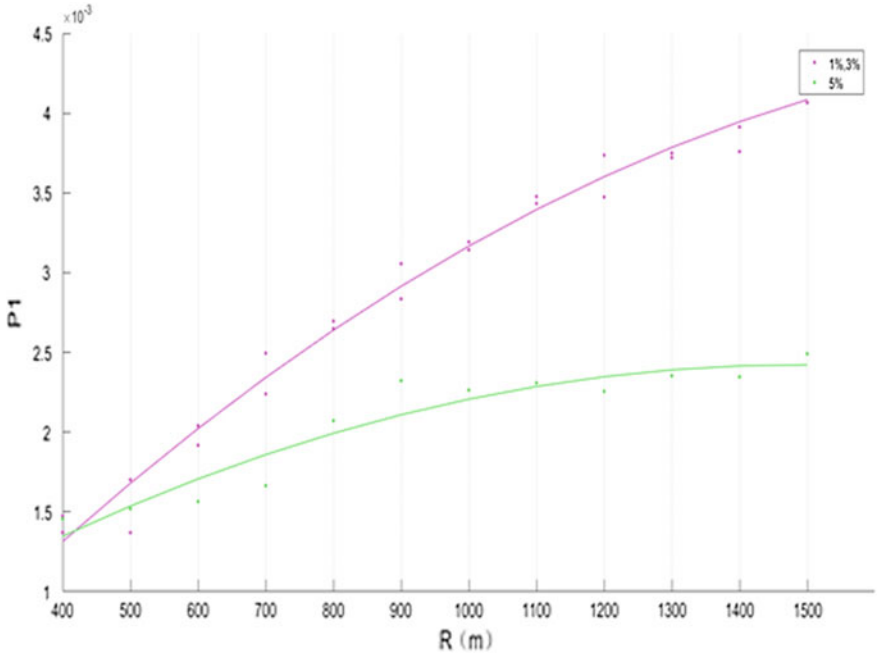


Fig. 11 Uphill and left-turn (fitting)

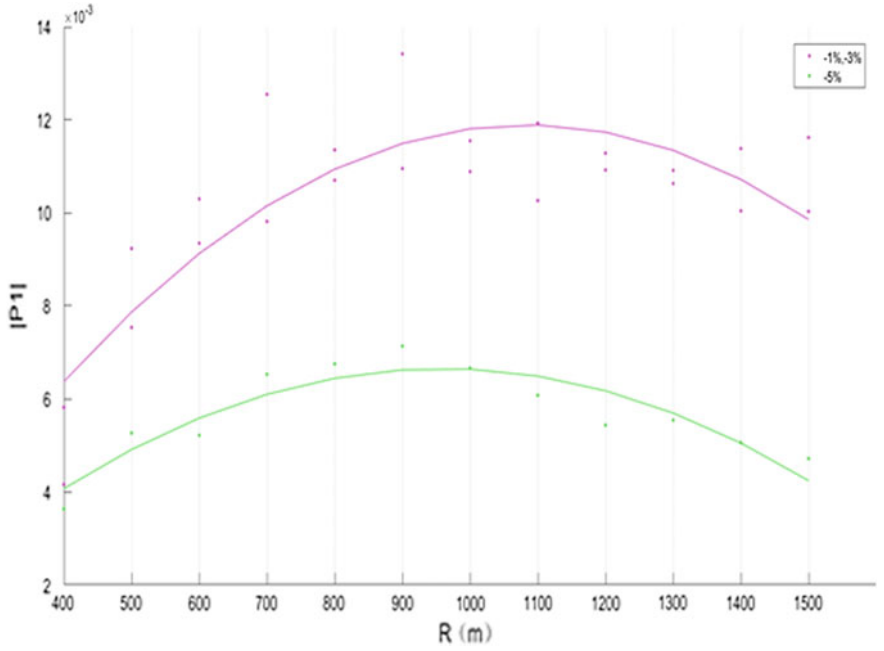


Fig. 12 Downhill and right-turn

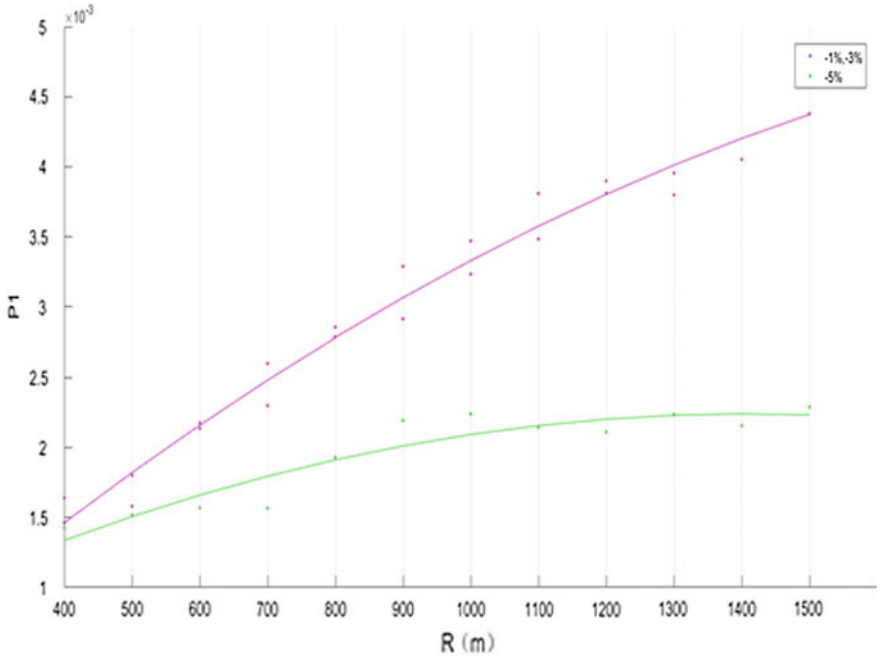


Fig. 13 Uphill and left-turn (fitting)

As the radius of the curve increases, $|P_1|$ increases, indicating that the radius of the left-turn curve increases resulting in a trajectory that tends to be flat (see Figs. 10 and 13).

4.2.2 Effect of Slope on the Track

When the curve direction is consistent, the effects of slope type (uphill or downhill) on P_1 are not much different (see Figs. 14 and 15).

When the slope is sharp ($\geq 5\%$ or $\leq -5\%$), $|P_1|$ will be significantly smaller, and the corner of the vehicle will be significantly larger than the curve with a low grade slope. As the radius increases, the difference increases. Therefore, the effects of the slope on the trajectory will more obvious as the radius increases. When the vehicle is driving on a small radius left-turn curve, the slope has little effect on its trajectory.

4.2.3 Effect of Curve Direction on the Track

When the vehicle is on the right-turn curve, the $|P_1|$ value is much larger than when it is on the left-turn curve (see Figs. 14 and 15). The vehicle's trajectory is more gradual

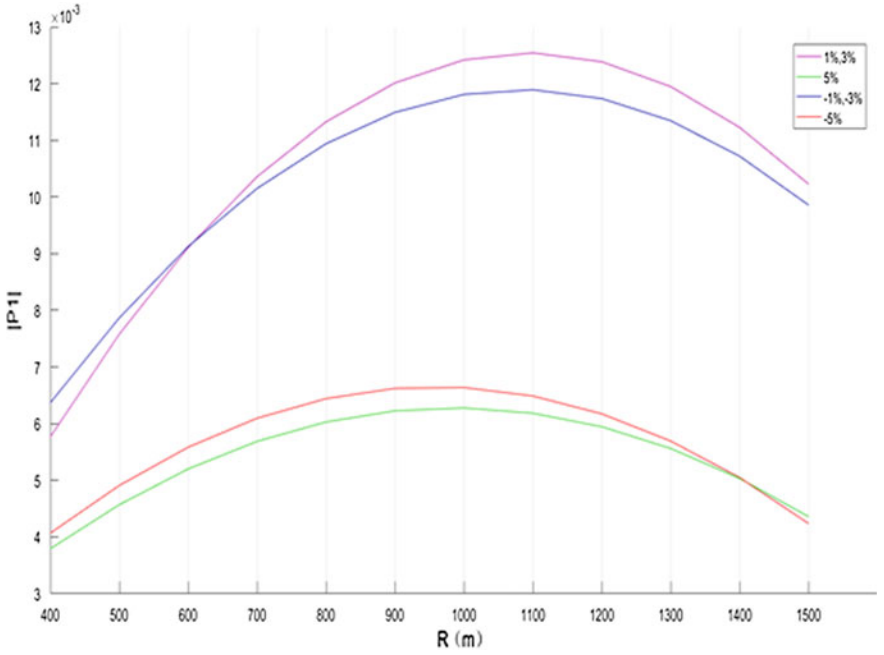


Fig. 14 Right-turn

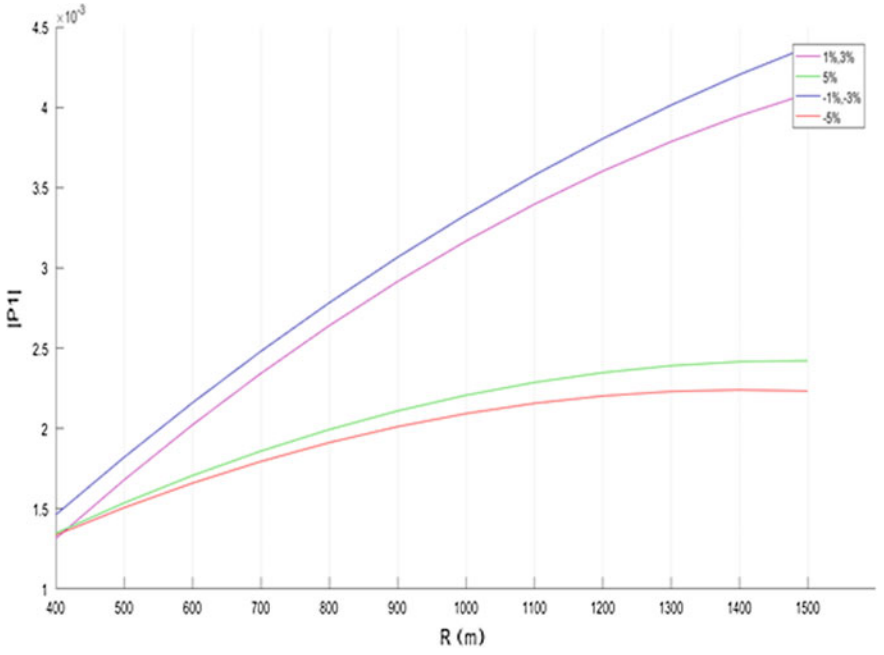


Fig. 15 Left-turn

when turning right and the driver is more inclined to pass the right-turn curve in a straight line.

5 Discussion

When the two research analyses are combined, the influence of the combined alignments on lane departure can be accurately revealed. As can be seen, the larger the radius of the curve, the smaller the lane departure, and the trajectory of the vehicle will be more gradual. When the curve radius exceeds 1000 m, the lane departure may increase due to the difficulty of the driver's perception of the curvature of the road. This shift is obvious on the right-turn curve, where the driver will go through the curve with a larger turning radius. When the curve radius is 1500 and 400 m, the lane departure and $|P_1|$ value are approximately same, which means that not only does the sharp turn have a negative impact on driving, but the curves with large radii are also be detrimental to driving safety. JTG D20-200 [6] specifies that the minimum radius of a circular curve is 400 m when the design speed is 100 km/h, and the maximum radius does not exceed 10,000 m. The theoretical maximum radius is obviously too big from a lane departure perspective.

The greater the slope of the curved road, the more likely the vehicle will be offset and the greater the degree of lane departure. When the gradient is -3 to 3%, the slope has little effect on the turning trajectory. When the slope is $\geq 5\%$ or $\leq -5\%$, the difficulty in controlling the direction of the car is obviously increased, and the trajectory is more prominent. As the radius increases, the trajectory will be more affected by the slope. This means that a sharp slope is not conducive to the driver determining the shape of the road, and it is more likely to cause a lane departure. In addition, the type of slope (uphill or downhill) has no significant difference in driving behavior.

The lane departure and trajectory when turning right are generally less affected by the slope than when turning left (see Figs. 14 and 15). The lane departure when the vehicle turns right is relatively large relative to the left-turn, but the trajectory is more gradual. It can be considered that since the driver is located on the side of the vehicle, the difference in the curve direction affects the driver's behavior. In China, a vehicle is required to drive on the right, the driver's seat is located on the left side of the vehicle, and the driver is far away from the right rear view mirror [26–28]. This is not conducive to drivers accurately observing the change in curvature of the current road section when turning right, and may result in the driver tending to go through the curve with a gentler trajectory when turning right, which will cause a larger lane departure. Therefore, considering the influence of the driver's seat position on the driver's driving process, different linear design specifications should be formulated according to the difference in the curve direction to ensure safety during turning.

6 Conclusions

The purpose of this study was to design safer combined alignments on highways by studying the impact of the horizontal and vertical combined route on lane departure. In this study, a multi-curve expressway was created using a driving simulator, and the lane departure and trajectory of each curve were analyzed separately including the distance between the vehicle centerline and the lane centerline, and the change law of |P1|, which can better explain the impact of the combined alignments on lane departure [29–31].

According to the results, the curve radius, slope and curve direction influence lane departure. Specifically, lane departures increase as the horizontal curvatures and slopes increase. However, it is worth noting that there is more risk to turning right than left. The right-turn curve is not conducive to the driver's observation of the road alignment. When the curve radius exceeds 1000 m, lane departure will start to increase. The lane departure on a right-turn will be greater than that on a left-turn curve. The radius of the vehicle trajectory on a right-turn curve will be larger than the road radii. In addition, the slope is positively correlated with lane departure.

In this paper, the safety of a vehicle is explored given the combined alignments of an expressway using the influence of curve radius, slope and curve direction on the lane departure. Future work should focus on the impact of the curve direction on the driver's behavior and formulate corresponding design specifications.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 61503007); Nova Cooperation Program of Beijing (Z171100001117131); Youth Top Talent Training Program of Beijing Education Committee and General Program of science and technology plan of Beijing Education Committee (SQKM201810009007); Nova Cooperation Program of Beijing (Z171100001117131).

References

1. Eustace D, Almuntairi O, Hovey PW, Shoup G (2014) Using decision tree modeling to analyze factors contributing to injury and fatality of run-off-road crashes in Ohio. In: 93rd Annual meeting transportation research board, Washington, DC
2. Lord D, Brewer MA, Fitzpatrick K, Geedipally SR, Peng Y (2011) Analysis of roadway departure crashes on two-lane rural roads in Texas. Federal Highway Administration. Rep, pp 3–5
3. Liu C, Subramanian R (2009) Factor related to fatal single-vehicle run-off-road crashes. Gender
4. Chen Y, Quddus M, Wang X (2018) Impact of combined alignments on lane departure: a simulator study for mountainous freeways. Transp Res Part C Emerg Technol 86:346–359
5. Ministry of Transport of the People's Republic of China (2006) Design specification for highway alignment. JTG D20—2006. China Communications Press
6. Torbic DJ, Harwood DW, Gilmore DK, Pfefer R, Neuman TR, Slack KL, Hardy KK (2004) Guidance for implementation of the AASHTO strategic highway safety plan—a guide for reducing collisions on horizontal curves. Transportation Research Board of the National Academies, Washington, DC

7. Wilson TD (1968) Road safety by design. Highway Eng 15:23–33
8. Park BJ, Fitzpatrick K, Lord D (2010) Evaluating the effects of freeway design elements on safety. Transp Res Rec J Transp Res Board 2195(1):58–69
9. Jalayer M, Zhou H (2017) Exploratory analysis of run-off-road crash patterns. In: 96th Annual meeting transportation research board. National Academies, Washington, DC
10. Miaou SP, Lum H (1993) Statistical evaluation of the effects of highway geometric design on truck accident involvements. Transp. Res. Rec. J. Transp. Res. Board 1407:11–24
11. American Association of State Highway Transportation Officials (AASHTO) (2011) A policy on geometric design of highways and streets. Association of State Highway Transportation Officials, Washington DC
12. Fatality Reporting Analysis System (2011) National highway traffic safety administration. www-fars.nhtsa.dot.gov/Main/index.aspx. Accessed 21 June 2018
13. Cartes O (2002) Multiple perspectives. In: Fuller R, Santos JA (eds) Human factors for highway engineers. Elsevier Science Ltd., pp 11–21
14. Lamm R, Psarianos B, Mailaender T, Choueiri EM, Heger R, Steyer R (1991) Highway design and traffic safety engineering handbook. USA, McGraw-Hill, Professional Book Group, New York, NY
15. Glennon JC, Weaver GD (1971) The relationship of vehicle paths to highway curve design. Photography
16. Good MC (1978) A review of empirical studies of driver-vehicle behavior on road curves, road curve geometry, and driver behavior. Australian Road Research Board
17. Abdel-Aty M, Pemmanaboina R, Hsia L (2006) Assessing crash occurrence on urban freeway by applying a system of interrelated equations. Transp Res Rec J Transp Res Board 1953:1–9
18. Hanno D (2004) Effect of the combination of horizontal and vertical alignments on road safety. The University of British Columbia
19. Cafiso S, Di Graziano A, Cava GL (2005) Actual driving data analysis for design consistency evaluation. Transp Res Rec J Transp Res Board 1912:19–30
20. Wang X, Wang T, Tarko A, Tremont PJ (2015) The influence of combined alignments on lateral acceleration on mountainous freeways: a driving simulator study. Accid Anal Prev 76(3):110–117
21. Wu X, Wang X, Lin H, He Y, Yang L (2013) Evaluating alignment consistency for mountainous expressway in design stage: a driving simulator-based approach. In: 92nd Annual meeting transportation research board. National Academies, Washington, DC
22. Lin Y, Niu J, Ying X (2011) Study on characteristics of vehicle path in curves on two-lane highways. J. Highw. Transp. Res. Dev. 28(3):113–117
23. Said D, Hassan Y, Halim AE (2007) Quantification and utilization of driver path in improving the design of highway horizontal curves. In: The 86th Transportation Research Board Annual Meeting, Washington DC
24. Spacek P (2005) Track behavior in curve areas: attempt at typology. J. Transp. Eng. 131(9):669–676
25. Yang Q, Overton R, Han LD, Yan X, Richards SH (2011) Driver behaviors on rural highways with and without curbs: a driving simulator-based study. In: The 90th Transportation Research Board Annual Meeting, Washington DC
26. Zhang RH, He ZC, Wang HW, You F, Li KN (2017) Study on self-tuning tyre friction control for developing main-servo loop integrated chassis control system. IEEE Access 5:6649–6660
27. Sun XJ, Zhang H, Meng WJ, Zhang RH, Li KL, Peng T (2018) Primary resonance analysis and vibration suppression for the harmonically excited nonlinear suspension system using a pair of symmetric viscoelastic buffers. Nonlinear Dyn 94:1–23. https://doi.org/10.1007/s11071-018-4421-9
28. Xiong H, Zhu X, Zhang R (2018) Energy recovery strategy numerical simulation for dual axle drive pure electric vehicle based on motor loss model and big data calculation. Complexity, pp 1–16
29. Kong X, Xia F, Ning Z, Rahim A, Cai Y, Gao Z, Ma J (2018) Mobility dataset generation for vehicular social networks based on floating car data. IEEE Trans Veh Technol 67(5):3874–3886

30. Rahim A, Kong X, Xia F, Ning Z, Ullah N, Wang J, Das S (2018) Vehicular social networks: a survey. Pervasive Mob Comput 43:96–113
31. Kong X, Song X, Xia F, Guo H, Wang J, Tolba A (2018) LoTAD: long-term traffic anomaly detection based on crowdsourced bus trajectory data. World Wide Web 21(3):825–847

Research on Driving Fatigue Level Using ECG Signal from Smart Bracelet

Jiyuan Tan, Xiang-yun Shi and Weiwei Guo

Abstract Fatigue level was studied using ECG signals collected from smart bracelet, and the data collected from the smart bracelet could reduce the interference to the driver during the experiment, increase the accuracy of the data. In order to accurately fatigue level, the algorithm of driving fatigue level was proposed based on multi-index fusion theory. Aiming at the missing status of the comprehensive indicator T value from principal component analysis, BP neural network was used for data compensation. The threshold of fatigue level was determined on the basis of the peak trough of the comprehensive index T value. Based on experimental analysis, it was found that the algorithm can effectively identify the wide awake and severe fatigue states when the data was missing.

Keywords Smart bracelet · BP neural network · Comprehensive index T · Fatigue level

1 Introduction

With the number of motor vehicles sustained growth in China, road traffic safety problems were becoming more and more serious, road traffic accidents gradually became one of the main reasons causing human casualties, 57% of catastrophic accidents related to driver driving fatigue, driving fatigue is one of the main factors causing major traffic accidents [1, 2]. Driving fatigue refers to the phenomenon that the driving skills to decline due to continuous driving for a long time. Therefore, strengthening the research of driving fatigue detection technology to preventing the occurrence of driving fatigue is of great significance for improving road traffic safety.

At present, the main methods for studying driving fatigue at home and abroad include driver detection and objective detection [3]. Driver detection was susceptible

J. Tan · X. Shi · W. Guo (✉)
Beijing Key Lab of Urban Intelligent Traffic Control Technology,
North China University of Technology, Beijing 100144, China
e-mail: guoweiwei@ncut.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_62

to factors such as gender and age. During the experiment the driver had intentionally concealed feelings or bias conditions to affect the accuracy of results [4]. The accuracy of EEG and ECG signal detection was high in objective detection, but the detecting equipment was expensive [5]. Moreover, the detecting equipment interferes with the driver. Therefore, this paper collected ECG signals from smart bracelets for driving fatigue level research. The Electrocardiogram (ECG) reflects the activity of the sympathetic and parasympathetic nerves [6]. It reflects the fatigue state of the driver to some extent. Rongrong et al. were using fast independent component analysis (FastICA) and digital filter to process the original signals, the discriminant criterion of fatigue was obtained from the training samples by using Mahalanobis distance, and then the average classification accuracy was given by 10-fold cross-validation. This method can give well performance in distinguishing the normal state and fatigue state [7]. Wu et al. collect ECG data under normal state and fatigue state and analyze the collected samples by using Kernel Principal Component method, Analysis results show that this method can effectively distinguish between normal state and fatigue state [8]. Currently, there are few studies on driving fatigue level at home and abroad. There are more consideration of fatigue and non-fatigue judgments. The detailed fatigue level studies were lack of and level missing state of data was less consideration in fatigue analysis. Based on the above problems, this paper proposed a fatigue level algorithm to conduct a more detailed level study on driving fatigue.

2 ECG Data Collection

In this paper, driving fatigue experiment is carried out with driving simulator, which provided data support for the algorithm of fatigue level. Firstly, a simulation scene of highway was designed. This scene is easy to cause fatigue for the driver. Then the experimental process and precautions were designed. Finally, the main equipment of ECG data collection was introduced.

2.1 Experimental Scene Design

As shown in Fig. 1, this paper used UC-winroad software to build a road network consisting of two 3/4 circles with a radius of 8 km and two straight segments of 14 km, speed limit of 100 km/h, two-way six lanes. The scene both sides of the road was designed monotonous to get tired easy for the driver. The simulated driving environment was sunny. Driving simulated laboratory was a white light with moderate brightness.

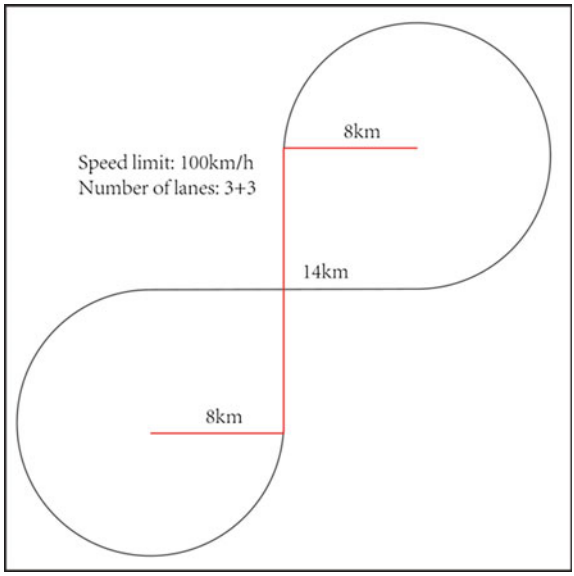


Fig. 1 Experimental scene setting

2.2 Experimental Process and Precautions

In this experiment, a total of nine graduate students are recruited. The age of the subjects was 22–27 years old, driving age was more than 1 year, without any history of disease, visual function was normal, skilled driving skills and strong safety awareness. The experiment time was 3 h arranged at 12:00–15:00 and 18:00–21:00. The bracelet were wore on subjects and the video capture device was placed before the experiment. Then subjects drove the driving simulator on the designed highway. At the same time the data required for the experiment was collected. The specific process was as follows:

Step 1: Filling in the relevant questionnaires;
Step 2: Opening the experimental scene, adjusting the driving posture and perspective, and informing the driver about the experimental rules and precautions;
Step 3: Placing the camera to clearly capture the face expression of the driver and the upper body movements;
Step 4: wearing the bracelet and closing to the wrist skin of the driver, connecting the app from the Bluetooth to test the stability of data transmission;
Step 5: Start the experiment;
Step 6: Filling in the relevant questionnaire.

Note: Keeping enough sleep at the night before the experiment, and do not drink stimulating beverages containing coffee for 3 h before the experiment; Simulating the actual driving environment during the experiment, trying to minimize unnecessary conversations and don't check your phone.

2.3 *Experimental Data Acquisition Device*

The cost of ECG signal acquisition equipment on the market was large, and it was easy to cause greater interference to the drivers. Therefore, this experiment get ECG data collection from a self-designed bracelet.

Comparing to other equipment, the bracelet was less expensive to manufacture and had less interference to the driver, allowing the driver to concentrate on driving. The functions that the bracelet needed to implement were shown in Fig. 2: Adjusting the bracelet acquisition mode to perform mode A, and collecting the ECG raw data at a sampling interval of 40 ms. The raw ECG signals were processed to obtain real-time heart rate and RR interval values, and the results were added to ble_tx_fifo. Long press the touch button to boot, the display could be switched from clicking the button, the ECG signal acquisition mode could also be switched from pressing the button.

3 Algorithm of Fatigue Level

As shown in Fig. 3, the RR intervals of equal time intervals were obtained after linear interpolation, and the time-frequency domain indicator was calculated from the ECG signal. After the principal component analysis method [9, 10], the dimension was reduced, and the comprehensive index T was obtained after fusion. The T value of the missing data was performed to compensate by Then BP neural network [11, 12]. Finally, the driving fatigue was divided into wide awake, mild, moderate and severe fatigue in the light of the Stanford Sleepiness Scale, and the threshold was calculated according to the corresponding formula.

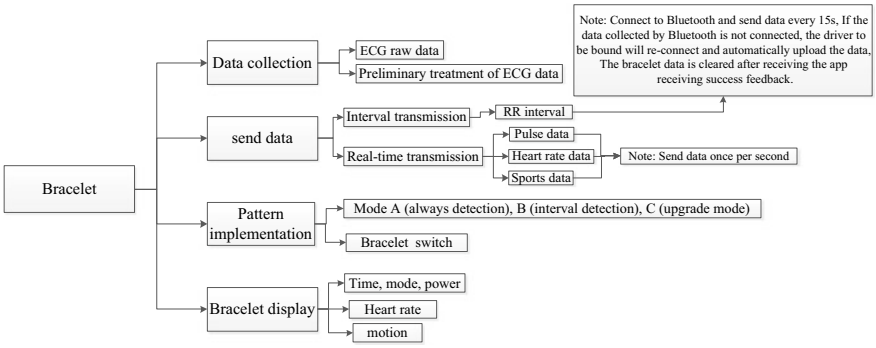


Fig. 2 Required functions of the bracelet

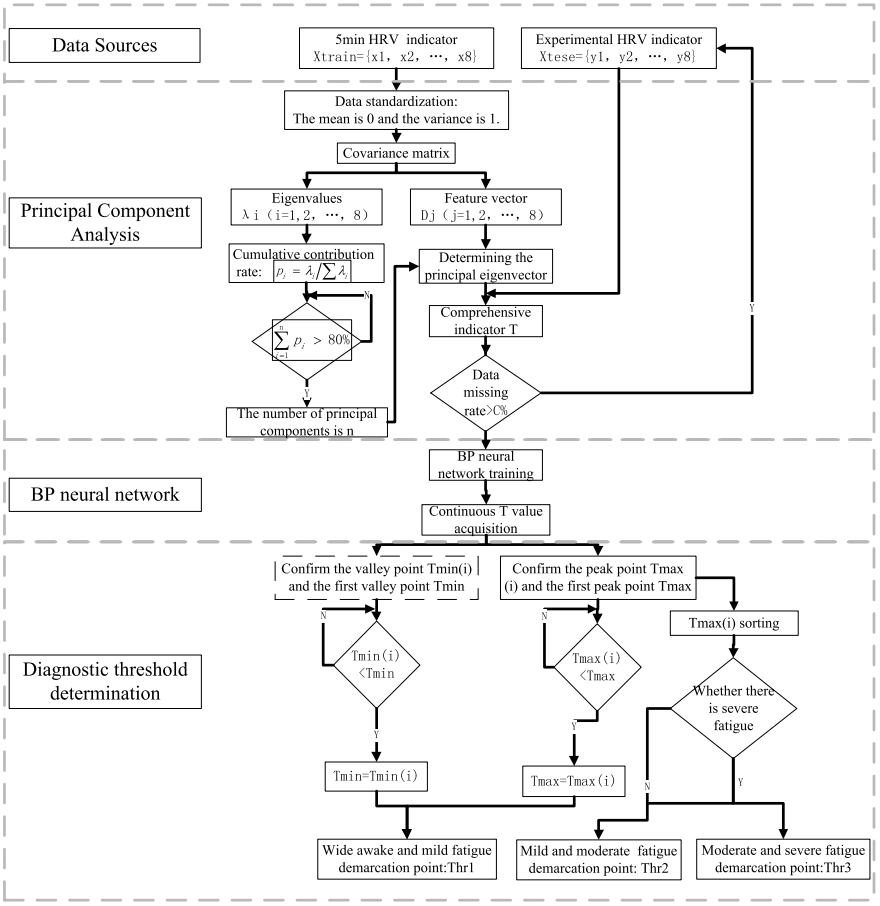


Fig. 3 Driving fatigue level flow chart

3.1 Comprehensive Index T Calculation

This paper selected commonly ECG time-frequency domain indicators AVER_RR, SDNN, RMSSD, PNN50, VARIANCE, LF, HF, LF/HF [13–15] as input variable of comprehensive index T. The RR interval of equal time interval was obtained from preliminary processing, and the physiological index values were calculated from the RR interval of the equal time interval. Subsequently, the ECG indicator vector X = [X1, X2, …, X8] was established. Principal component analysis was used to determine the principal elements and the contribution rate of each indicator, as shown in Fig. 4.

After obtaining the contribution rate of each indicator, the multivariate fusion was used to obtain the comprehensive index T, and the formula for calculating T was:

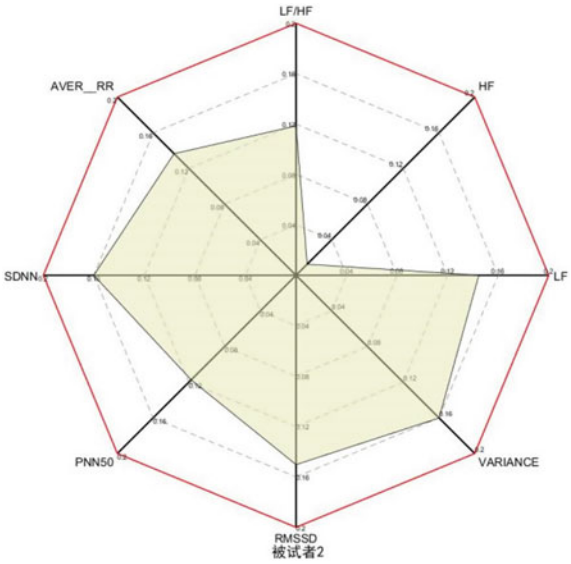


Fig. 4 Driver's contribution rate

$$T = \sum_{r=1}^{R} \frac{t_r^2}{\sigma_{t_r}^2} = x_i^T V_R \sum_{R}^{-2} V_R^T x_i \tag{1}$$

Among, σ_{t_r} was the standard error of the rth scale, T was the distance from the sample to the origin in the principal component analysis subspace.

As shown in Fig. 5, comprehensive indicators $\check{T}$ had data missing status. Therefore, the indicator T should be compensated.

3.2 Comprehensive Indicator T Compensation

In this paper, the feedforward BP neural network was used to compensation the T value and implemented from MATLAB programming. The format was:

net = newff(pn, tn, [n1 n3 n2], 'tansig', 'purelin', 'traincgf');

Among, net was BP neural network, Pn was the input matrix, tn was the output matrix, [n1 n3 n2] was the number of neurons in the network input layer, hidden layer and output layer, tansig was the transfer function of the hidden layer, purelin was the transfer function of the output layer, traincgf was the training function.

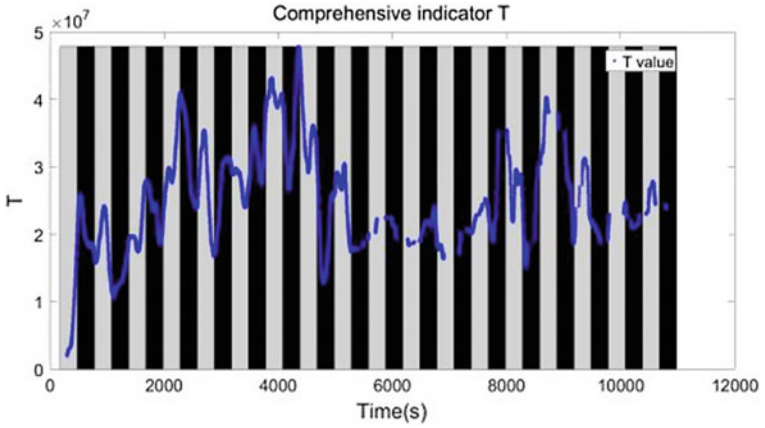


Fig. 5 Comprehensive indicator T

The maximum number of training sessions of BP neural network was 5000, the learning rate was 0.01, the training target error was set to 0.1×10^{-3}, and the number of hidden layers was determined from the following formula:

$$n3 = \sqrt{n1 + n2} + a \tag{2}$$

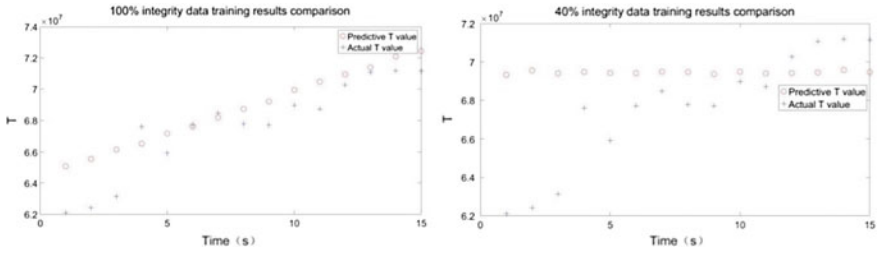
Among, a was a constant, the range is 5–10.

The data with complete T value was selected all to perform data loss and analysis, the training error results were shown in Table 1.

According to Fig. 6 and Table 1, when the data loss rate reaches 40%, the BP neural network could not be compensated more accurately. When the loss rate of T value data was 30%, BP neural network training was performed, the training results were shown in Fig. 7. It was obvious that the compensating results were in line with the data fluctuation trend.

Table 1 BP neural network training error

Data integrity (%)	Training error
100	4.01×10^{-4}
90	4.08×10^{-4}
80	4.24×10^{-4}
70	4.44×10^{-4}
60	8.22×10^{-4}



(a) 100% integrity data (b) 40% integrity data

Fig. 6 BP neural network training results

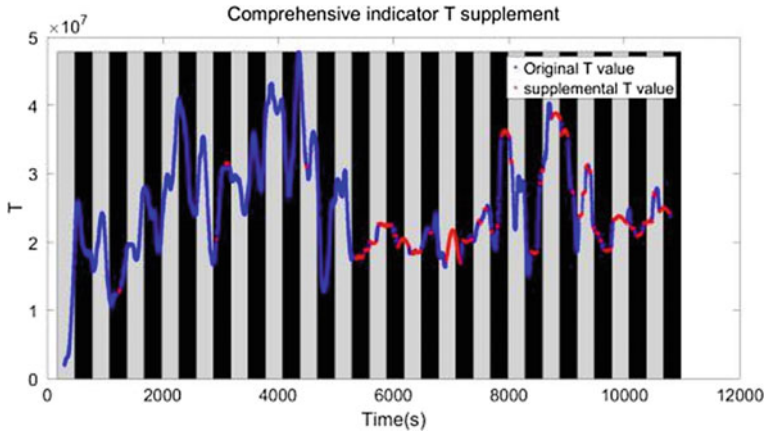


Fig. 7 BP neural network training results

3.3 Fatigue Level Threshold Determination

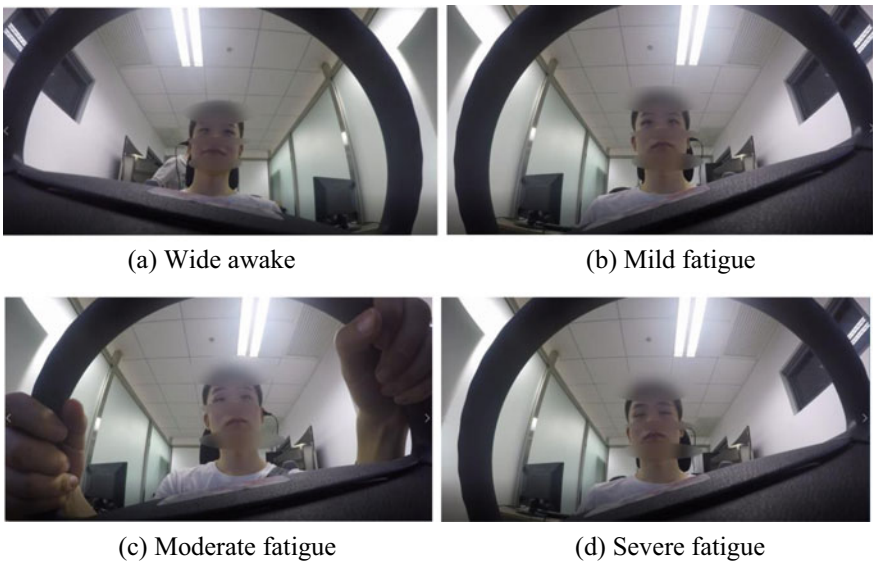
By analyzing the T-value data, it was found that the driving fatigue was fluctuating, which was consistent with the actual driving situation. Therefore, the T-value peak-to-valley point was extracted as a standard to determine threshold value of the fatigue level The driving state of the drivers was divided according to the Stanford Sleepiness Scale. The classification criteria are shown in Table 2.

Then based on the video data and the Stanford Sleepiness Scale the fatigue level was calculated every 5 min to used as a comparison the accuracy of driving fatigue level. The standard of the fatigue level was the proportion of the fatigue state of the driver within 5 min. If proportion of the wide awake state is the highest, the driver's driving state was wide awake within 5 min. An example of fatigue level was shown in Fig. 8.

5 min was selected as a time period for Fatigue level. The experimental time 3H was divided into 36 segments. Fatigue level divided into wide awake ($T(i) \leq Thr1$),

Table 2 Stanford sleepiness scale [16]

Fatigue level	Level	Status
Level 1: Wide awake	1	Feeling active, vital, alert, or wide awake
Level 2: Mild fatigue	2	Functioning at high levels, but not at peak, able to concentrate
	3	Awake, but relaxed; responsive but not fully alert
Level 3: Moderate fatigue	4	Somewhat foggy, let down
	5	Foggy, losing interest in remaining awake, slowed down
Level 4: Severe fatigue	6	Sleepy, woozy, fighting sleep, prefer to lie down
	7	No longer fighting sleep, sleep onset soon, having dream-like thoughts
X	X	Sleeping



(a) Wide awake (b) Mild fatigue

(c) Moderate fatigue (d) Severe fatigue

Fig. 8 Schematic diagram of fatigue level

mild fatigue (Thr1 < T(i) ≤ Thr2), moderate fatigue (Thr2 < T(i) ≤ Thr3), and severe fatigue (Thr3 < T(i)).

The comprehensive index T was obtained from BP neural network. After finding peaks and troughs in T value, according to the following formula the thresholds of the four fatigue states was calculated.

Wide awake and mild fatigue level threshold Thr1:

$$\text{Thr1} = \text{C1} * (\text{Tmin} + \text{Tmax}) \tag{3}$$

Mild and moderate level threshold Thr2:

$$\text{Thr2} = \text{Tmax}(\text{C2} * \text{length}(\text{Tmax(i)})) \tag{4}$$

Moderate and severe level threshold Thr3:

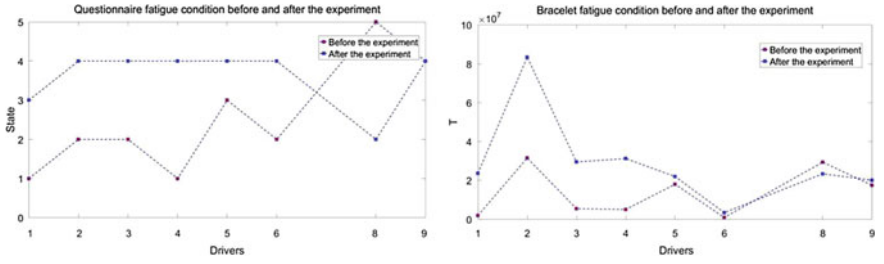
$$\text{Thr3} = \text{Tmax}(C3 * \text{length}(\text{Tmax(i)})) \tag{5}$$

Among, Tmin was the minimum valley point T value, Tmax was the minimum peak point T value, Tmax(i) was the peak point T value, Tmax(i) was the valley point T value, length was the T value point number, C1, C2 and C3 were level coefficients.

4 Experimental Verification

In the experiment, the data of the driver 7 was lost, so the fatigue level analysis was performed only for the data of the remaining 8 drivers. The fatigue level of the questionnaire was divided into 1–5; The bigger the value, the more fatigue. The results of the questionnaire were shown in Fig. 9a. Comparing with the data of the bracelet, as shown in Fig. 9b, except for the driver 9, the results of the other drivers were consistent with the results of the questionnaire. By observing the video data of the driver 9 it is found that the fatigue in the experiment was significantly higher than that before the experiment, but the driver 9 was always in a state of high fatigue, and there was a certain deviation from his own feelings. The results of the bracelet data could accurately reflect the fatigue of the drivers before and after the experiment.

Then, 8 drivers were drivered to fatigue level, and the fatigue level results were shown in Table 3. It could be concluded from the table that there was individual difference between the drivers, and it was impossible to quantitatively level all the drivers. The algorithm in this paper could effectively distinguish between wide awake and severe fatigue, but the level threshold for mild and severe fatigue was less recognition accuracy.



(a) Questionnaire (b) Bracelet

Fig. 9 Fatigue state before and after the experiment

Table 3 Fatigue level results of the drivers

Driv–er	Wide awake (%)	Mild fatigue (%)	Moderate fatigue (%)	Severe fatigue (%)	Comprehensive (%)	Thr1	Thr2	Thr3
1	100	100	75.86	–	80.56	1.14E+07	2.18E+07	–
2	100	50.00	95.21	–	80.56	4.10E+07	6.00E+07	–
3	88.89	100	–	–	97.22	1.57E+07	–	–
4	100	30.77	89.47	–	69.44	8.53E+06	1.84E+07	–
5	100	68.18	54.55	–	63.89	1.17E+07	2.93E+07	–
6	100	33.33	95.65	100	80.56	7.17E+05	4.58E+06	2.82E+07
8	100	80.00	62.29	100	75.76	1.75E+07	3.52E+07	8.36E+07
9	87.50	69.23	36.36	100	61.76	2.18E+07	4.02E+07	8.36E+07

5 Conclusion

In this paper, ECG data collection from the bracelet could effectively reduce the experimental interference for the driver and increase the accuracy of the data. Aiming at the ECG data, the algorithm of driving fatigue level was proposed based on principal component analysis and BP neural network.

Acknowledgements Supported by National Natural Science Foundation of China (Grant No. 61503005 & 61503007), Youth Top Talent Training Program of Beijing Education Committee (CIT&TCD201904013, PXM2019_0142) and General Program of science and technology plan of Beijing Education Committee (SQKM201810009007).

References

1. Gander PH, Marshall NS, James I et al (2006) Investigating driver fatigue in truck crashes: trial of a systematic methodology. Transp Res Part F Traffic Psychol Behav 9(1):65–76
2. Wei C (2000) Driving fatigue: prevention of french strafety. Soc Dev 10:43–44 (in Chinese)
3. Larue GS, Rakotonirainy A, Pettitt AN (2011) Driving performance impairments due to hypovigilance on monotonous roads. Accid Anal Prev 43(6):2037–2046
4. Declerck CH, Boone C, Brabander BD (2006) On feeling in control: a biological theory for individual differences in control perception. Brain Cognit 62(2):143–176
5. Jia-Jun Z, Hui-Zhang S, Yi-Qing W (2015) Design and implementation of multi-channel portable equipment for ECG and EEG acquisition. Chin J Med Phys
6. Billman GE (2011) Heart rate variability-a historical perspective. Front Physiol 2:86
7. Rongrong F, Hong W (2014) Detection of driving fatigue by using noncontact EMG and ECG signals measurement system. Int J Neural Syst 24(3):145006
8. Wu Q, Zhao Y, Bi X (2013) Driving fatigue classified analysis based on ECG signal. In: Fifth international symposium on computational intelligence & design. IEEE
9. Ting PH, Hwang JR, Doong JL et al (2008) Driver fatigue and highway driving: a simulator study. Physiol Behav 94(3):448–453
10. Chen LL, Zhao Y, Ye PF et al (2017) Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. Expert Syst Appl S0957417417300477

11. Ying Y, Jing S, Wei Z (2007) The monitoring method of driver's fatigue based on neural network. In: 2007 International Conference on Mechatronics and Automation. ICMA 2007. IEEE
12. Zizheng G, Yonggang T, Guozhong MA et al (2014) Recognition method of driving mental fatigue based on BP neural network. J Harbin Inst Technol 46(8):118–121
13. Jeong IC, Dong HL, Park SW et al (2007) Automobile driver's stress index provision system that utilizes electrocardiogram. In: Intelligent vehicles symposium. IEEE, pp 652–656
14. Takahashi I, Yokoyama K (2011) Development of a feedback stimulation for drowsy driver using heartbeat rhythms, vol 4, pp 4153–4158
15. Lanatà A, Valenza G, Greco A et al (2015) How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving. IEEE Trans Intell Transp Syst 16(3):1505–1517
16. Maclean AW, Fekken GC, Saskin P et al (2010) Psychometric evaluation of the Stanford sleepiness scale. J Sleep Res 1(1):35–39

An Analysis of the Travel Patterns of Pilgrimage Groups in Lhasa Tibet

Gang Cheng, Shu-zhi Zhao and Zong Wang

Abstract Religiously-influenced pilgrims are commonly seen in Lhasa, a political and cultural center in Tibet. This study observed the travel modes of pilgrims in Lhasa between 2011 and 2017. Information was gathered using questionnaires and interviews. Travel modes of pilgrims and non-pilgrims were compared to identify each group's regular patterns. The study assessed differences in the characteristics of the Pilgrims' travel behavior at different ages, and differences in the behavior of pilgrims with different employment statuses. Given the significant differences among pilgrims' backgrounds, surveys were conducted with subjects of different age groups and employment types. The study's conclusions represent the experience of local pilgrims. Transportation management departments can use this information to better understand the travel needs of pilgrims and provide a higher standard of travel services to ensure the smooth conduct of the pilgrimage. The results of this study also provide a reference for research about pilgrims in other areas, particularly in inhabited areas in Tibet. It can also provide quantitative data to support religious study in Tibet.

Keywords Pilgrims · Circumambulation for prayer · Travel behavior · Analysis of behaviors

1 Introduction

Circumambulation, which means walking all around a specific place, was a common practice among Tibetan Buddhists in Tibet and in other Tibetan areas in Sichuan, Yunnan, Qinghai, and Gansu. Walking along certain routes for prayer was a common

G. Cheng (✉)
College of Engineering, Tibet University, Lhasa 850000, China
e-mail: chenggang@utibet.edu.cn

G. Cheng · S. Zhao
College of Transportation, Jilin University, Changchun 130022, China

Z. Wang
College of Humanities, Tibet University, Lhasa 850000, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_63

religious activity, and remains an important custom in Lhasa. As time has passed, circumambulation for prayer has become a daily activity for common people [1], combining the building of new social relationships and the exercising of religious beliefs in the Qinghai-Tibet Plateau.

In this paper, pilgrimage refers to "circumambulation for prayer," which can be both an individual and group activity. In contrast to religious activities performed by monks in a monastery, this is a common practice in the daily lives of ordinary people, and provides both physical and spiritual benefits for believers.

More than 6,400,000 Tibetans live in the area covered by Tibet, Qinghai, Sichuan, Yunnan, Gansu, and other areas. Approximately 98% of these people live in China, with less than 2% scattered around the world. Most of them are devout Buddhists.

Lhasa is a sacred place in Tibetan Buddhism. The biggest dream of many Tibetans is to go on a pilgrimage to Lhasa, and live there full-time. This paper mainly refers to residents of the Chengguan District in Lhasa; it also includes residents from Deqing and Dazi, two suburban counties. The paper's authors have worked in Lhasa, facilitating the long-term observation and research of local resident lifestyles; intermittent observations have been collected since 2010. The travel patterns of pilgrims representing different ages and different social identities were observed, and interviews were conducted to gather more information. In 2015 and 2016, questionnaires were used to gather information from respondents for statistical analysis, and follow-up surveys were conducted for more information. Electronic/online questionnaires and paper questionnaires were both used due to respondent differences in education, language, and age. For people under 60, questionnaires were used for the survey; people over 60 were interviewed to maximize accuracy. Of 3000 questionnaires for pilgrims, 2686 were valid; there were 1236 valid questionnaires of a total 1500 administered to non-pilgrims.

2 Literature Review

The research on pilgrim travel behavior can be grouped into three subject areas. The first research area focuses on Tibetan pilgrimage. Karmay [2], studied pilgrimage rites in: "Arrows and spindles: a study of Tibet's history, mythology, rituals and beliefs." The scholar Chen integrated psychology to study the original intention, consciousness, and complex of the Tibetan religious believers. Zheng [3], Tian [4], and Fu [5] conducted comparative studies of pilgrimage and tourism, and analyzed the cultural phenomenon of pilgrimage. Zhang and Chen [6] studied the Tibetan Religious practice "circumambulation" and Personal Religious Experiences. Peng and Chen [7] studied Tibetan pilgrimage in the process of social change.

The second research area concerns public travel behavior. Both domestic and foreign scholars have studied travel behavior. Some have studied the elderly and different age groups [8]; others have studied gender, specifically focusing on women [9]; others have focused on poor rural areas, studying economic development. Other researchers have focused on a specific city (such as a historical and cultural city)

[10, 11] to study the travel behavior of specific populations. Many scholars have focused on families, assessing the quality and differences in behavioral data related to collective family activities [12, 13]. Mosa and Esawey [14] used a structural equation model and mixed Logit model to simulate family decision-making to travel to individual and collective activities. The results showed that family structure, social background, activity type, and other factors significantly impact the choice. Some scholars studied travel behavior using activity-travel chain theory to analyze the travel chain behavior of specific populations [15, 16] used an ordered probit regression model to study the activity and travel behavior of family members. Bueno et al. [17] studied the effects of transit benefits on employee travel behavior.

The third area of research involves pilgrim travel behavior. There have been few studies on the travel behavior of Tibetan pilgrims from the perspective of traffic travel; those few studies focus on pilgrim flow based on railway transportation or travel during the Religious Festival [18, 19].

In conclusion, studies on public travel behavior consistently show a connection between behavior and the local social development situation. China has a significantly different economic situation and social climate than other countries, so China serves as an effective model for research in Tibet. Economic conditions in Lhasa are relatively underdeveloped compared to the rest of China. The pilgrimage group studied here belongs to a unique group compared to other groups studied for their travel behavior. The trip at the center of this research is a special daily travel event with specific religious and social meaning. Previous studies about the Tibetan pilgrimage trip have mainly focused on qualitative factors. There have been few quantitative studies. This research examines pilgrimage groups with a strong religious background in Lhasa, using both qualitative and quantitative methods. This fills any missing quantitative analysis elements and provides data to support other Tibetology studies.

3 The Distinctiveness of the Pilgrimage Trip

Pilgrim travel is dynamic, with strong liquidity. The essence of the pilgrimage is to walk and pray around a particular route. In Tibetan language, it is expressed as: turning through. Given the specificity of this pilgrimage trip, pilgrim travel is divided into two parts: basic travel and travel driven by the Tibetan religious practice, "circumambulation." The basic travel is defined as the journey from the place of the pilgrims to the place of pilgrimage; the religious practice travel is a pilgrimage trip.

3.1 The Distinctiveness of Pilgrimage

In engaging in a pilgrimage, the pilgrim becomes proud of the long distance of the walk, the blessing ceremony, and the longer time spent than other people practicing

Tibetan customs [1]. The pilgrimage to the holy land is a religious social custom; it has a wide audience; and the pilgrims in Lhasa cover almost all ages.

3.2 The Distinctiveness of Circumambulation Trip

① The distinctiveness of transportation forms. Walking is a basic transportation mode for pilgrimages and is the core transportation mode for the circumambulation trip. The pilgrimage activity includes chanting, rotating the prayer wheel, kowtowing to complete the supplication, and the blessing.

② The distinctiveness of the trip scope. The pilgrimage activity area is relatively focused, with circumambulation occurring around Jokhang Monastery and the Potala Palace.

③ The distinctiveness of spatial form. The spatial form of the pilgrimage trip is greatly influenced by Tibetan Buddhism. In the Tibetan language, circumambulation also means "paying respect to Buddha," and coming closer to Buddha, as the highest center of sacredness. As history has unfolded and the city has developed, six circumambulation paths have formed around Jokhang Monastery and the Potala Palace: Nangkhor, Bakhor, Lingkhor, Tsekhor, Duikhor, and Mekhor. Based on their own preferences, pilgrims can choose to pray along any one of the six roads. Khor is Tibetan, and means circle. From a spatial perspective, the pilgrimage trip takes an oval shape. Fig. 1 shows a sketch of the area.

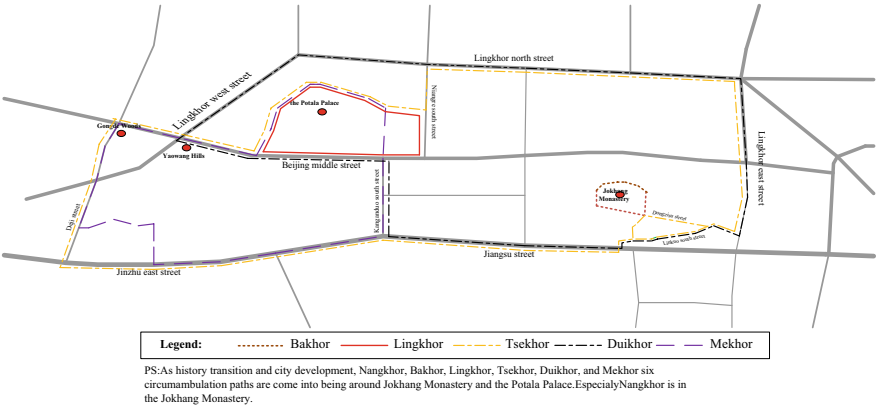


Fig. 1 Map of the circumambulation route

Table 1 Comparison of the travel ratio between pilgrims and non-pilgrims

Time node	Working day			Non-working day			Average travel ratio
Groups	Survey number	Travel ratio (time/day)	S.E	Survey number	Travel ratio (time/day)	S.E	
The pilgrims	2686	2.37	0.005	2686	2.13	0.005	2.25
The non-pilgrims	1336	2.43	0.012	1236	2.01	0.012	2.22

4 A Comparative Analysis on the Travel Characteristics of the Pilgrims and the Non-pilgrims

This study focused on people over 24-years-old, and analyzes the differences in travel behavior between pilgrims and non-pilgrims. Non-pilgrims are defined as people who live and work in the area, but don't complete the pilgrimage.

4.1 Travel Ratio

Based on the characteristics of people in Lhasa, working days and non-working days were selected as time nodes for the study. Sample data statistics (Table 1) show that the average travel ratio of pilgrims is 2.25 trips/day (number of trips per day). The average travel ratio of non-pilgrims is 2.22 trips/day; there figures are substantively the same. The pilgrims' travel ratio is less affected by time nodes; the difference in the pilgrims' travel ratio between the two time nodes was not as significant as with the non-pilgrims. Some non-pilgrims travel on working days because of work factors; however, on non-working days, they are not affected by work. As a result, travel rates decrease significantly on non-working days.

4.2 Travel Mode

There are two travel mode stages for pilgrims: the basic travel phase and the circumambulation for prayer (CFP), which is normally based on walking. Basic travel is the journey from the place of the pilgrims to the place of pilgrimage, the religious practice travel is a CFP trip. While in the basic travel phase, the pilgrims face the same travel choices as non-pilgrims. Table 2 describes the main travel modes of pilgrims and non-pilgrims.

The analysis revealed that the bus is the first transportation choice for pilgrims and non-pilgrims. Pilgrims depend the most on the bus, at 67%; walking is the second

Table 2 Comparison of main travel modes of pilgrims and the non-pilgrims (%)

Travel mode	Walking	Bicycle	Bus	Taxi	Car	Tricycle	Combined travel	Total
The pilgrims	11.7	0.2	67	3.7	11.1	3.8	2.5	100
The non-pilgrims	3.2	0.3	45	14	35	1.1	1.4	100

most common, at 11.7%. Therefore, the pilgrims' main travel modes are public transportation and walking, and other travel modes are supplemental.

4.3 Travel Purpose

Some travel purposes are shared between pilgrims and non-pilgrims, including working, studying, shopping, fitness, seeing a doctor, and visiting relatives and friends. The pilgrimage itself is the distinction of travel purpose between the pilgrims and the non-pilgrims.

4.4 Travel Time Distribution

The average daily travel time for pilgrims is 36 min; the average daily travel time for non-pilgrims is 26 min. This means that pilgrims travel of average of 10 min more than non-pilgrims per day. For both pilgrims and non-pilgrims, travel times of 21–40 min accounted for more than 50%. For pilgrims, travel time falling in the range of 0–20 min accounted for approximately 33%. For non-pilgrims, trips between 0 and 20 min make up approximately 25% of the total. Figure 2 shows the distribution of specific travel times.

4.5 Departure Time

There is a significant difference in the distribution of departure times from the residences of pilgrims and the non-pilgrims. Pilgrim departures are concentrated between 7:00–9:00. Peak travel is around approximately 8:00. Few pilgrims travel between 12:00–14:00; however, there is a small travel peak at 15:00. The number of the pilgrims then gradually decreases, with pilgrimage almost stopping after 20:00. The time range of non-pilgrim departures and travel is relatively even. Peak travel is around 8:00–9:00. There are other 2 small peaks at about 15:00 and 18:00; travel stops after 22:00. The morning rush hour for pilgrim travel is approximately one hour earlier than for non-pilgrims. Figure 3 shows the distribution of departure time.

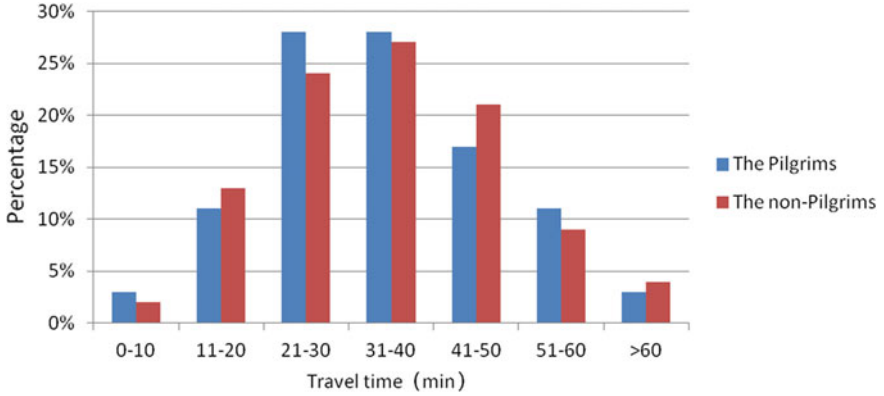


Fig. 2 Distribution of travel times for pilgrims and non pilgrims

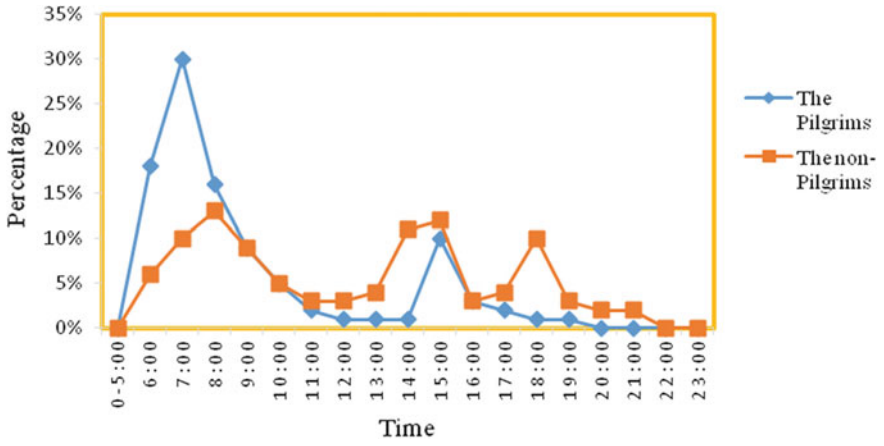


Fig. 3 Distribution of departure time for pilgrims and non-pilgrims

A comprehensive analysis of the travel ratio, travel mode, travel purpose, travel time, and pilgrim departure time reveals the following. There was no significant change in the travel ratio of pilgrims on working and non-working days. The average travel ratio is pilgrims was similar to non-pilgrims. For travel mode, pilgrims mainly rely on the bus for basic travel. Therefore, public travel time is relatively long in duration, and pilgrims and non-pilgrims are relatively close to each other in overall in travel time. Comprehensive analyzing travel time and travel distance suggests that urban traffic in Lhasa is relatively underdeveloped, and the effectiveness of traffic management needs to be strengthened. For example, the pilgrims' core activities area overlaps with the city's business district, resulting in traffic that is not very smooth. Further, the pace of life in the city is slow, and the hypoxia caused by high altitudes,

is not suitable for strenuous exercise. Religious factors have some impact on the pilgrimage; few people pilgrimage after 21:00.

5 Analysis of Travel Behavior Changes in Different Pilgrim Age Groups

To quantify the effects of different factors on pilgrim travel, study subjects were grouped by age. Table 3 provides the criteria used to describe the different groups.

5.1 Travel Rate Changes by Age

Figure 4 depicts the change in pilgrim travel ratio with age. As ages increase, the pilgrim travel ratio decreases. The most active travel ratio occurs during age stage 3, at 2.43 times/day in trips per day. The age stage 8 appears to be the most inactive stage, at 1.83 times/day.

The analysis indicates differences in travel ratios between the different age stages, and shows a travel rate that changes from high to low. The difference between age stage 4 and age stage 5 was highest, reaching to the maximum of 0.23. It is generally during age stage 5 that many people choose to enter a more leisurely lifestyle. As ages increase, the travel ratio gradually decreases.

Table 3 Age characteristics of Tibetan people in Lhasa

Age group	Age range unit: year	Typical identity	Typical life state	If consider it as a research object of the pilgrims
Age group 1	0–18	Student	Studying	No
Age group 2	19–23	Student, social man	Studying, working	No
Age group 3	24–48	Social man	Working	Yes
Age group 4	49–60	Social man	Retired/working	Yes
Age group 5	61–65	Social man	Retired	Yes
Age group 6	66–70	Social man	Retired	Yes
Age group 7	71–75	Social man	Retired	Yes
Age group 8	75 or above	Social man	Retired	Yes

Social man: A person with the economic ability to live in society

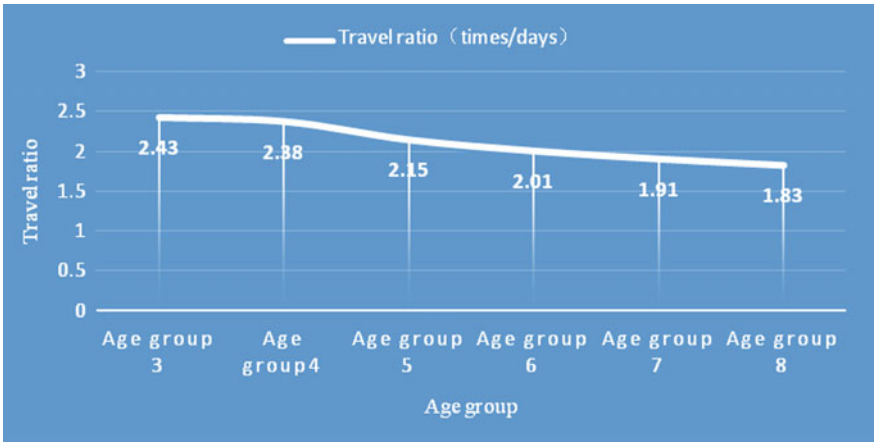


Fig. 4 Changes in pilgrim travel ratio with age

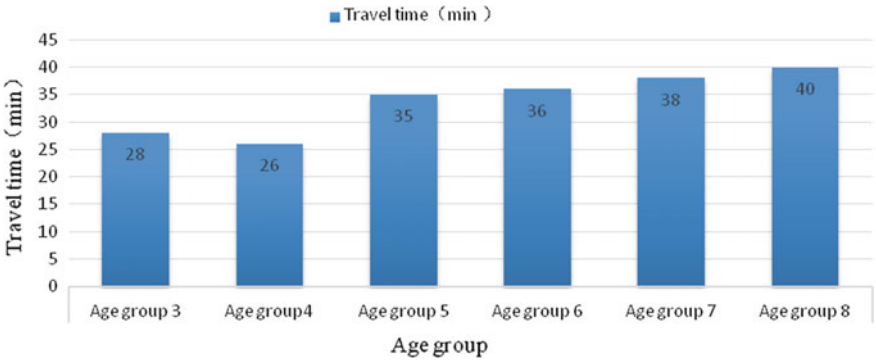


Fig. 5 Travel time of pilgrims with age

5.2 Travel Time Changes by Age

As Fig. 5 shows, age has little effect on average pilgrim travel time. The core area of the pilgrimage is a fixed place, and the distance from the starting point to the destination changes little as ages change. As pilgrims age, there is a slight increase in travel time due to physical aging and slower action, which results in them getting on and off more slowly.

5.3 Changes in Travel Distance with Age

The travel distance is mainly composed of two stages: basic travel and CFP walking.

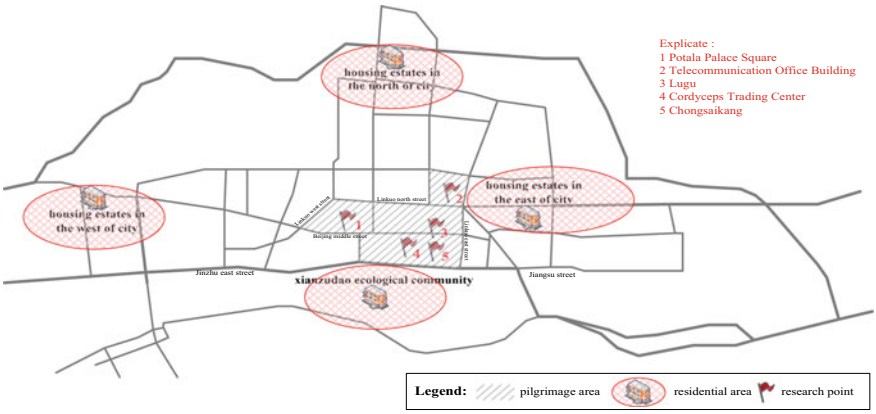


Fig. 6 Research groups based on location

① Basic travel distance. To study the basic travel distance, we selected four groups of pilgrims living in relatively concentrated areas in four directions (east, south, west and north), including housing estates in the north of city, housing estates in the east of city, housing estates in the west of city, and the Xianzudao ecological community (located in the south of the city). The core activities associated with pilgrim daily travel are mainly located in the central area of the city. To study the travel distance from the four residential areas to the central area, four points were selected as the starting points: the telecommunication office building, Chongsaikang, cordyceps trading centre, Lugu, Potala Palace Square served as the starting point of the pilgrimage, which was the endpoint of the basic travel (Fig. 6). Based on the survey, the basic travel distances for pilgrims is mainly concentrated in the 1.7–6.4 km range, with an average basic travel distance of 5.3 km.

② The CFP walking distance. Because of the varying lengths of circumambulation paths, the CFP walking distances are not all the same based on personal preferences. There are 6 circumambulation paths recognized in Lhasa, with lengths ranging between 1.5 and 8 km. Influenced by Tibetan Buddhism, the walking distance increases as the length of turning channels increases, on an odd number basis.

③ Travel distance. The average travel distance for the pilgrims is:

$$L_{CS} = l_{JC} + l_{ZJ} + \omega_{LJ} \tag{1}$$

l_{JC}: The distance of pilgrim travel from home to the starting point; the recommended value is 5.3 km; $l_{ZJ} = \alpha n\, \alpha$: The conventional distance to a Tibetan religious practice "circumambulation." Conventional values are 1.5, 2, 7, and 8 km; n: The number of Tibetan religious practice "circumambulations." Conventional values are

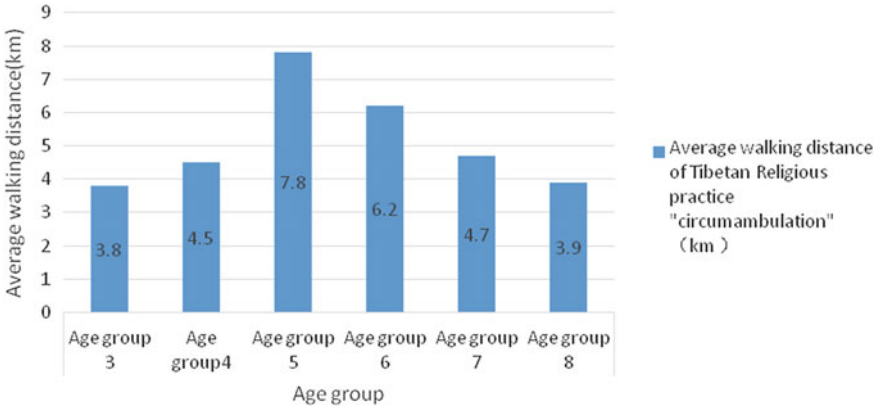


Fig. 7 Average walking distance for CFP across different age groups

1, 3, 5, 7, 9, and other odd numbers ω_{LJ}: Distance for connection of the circumambulation paths; the conventional value is 0.9 km; L_{CS}: The total distance of pilgrims; l_{ZJ}: The distance traveled by pilgrims in the course of "circumambulation."

Statistics show the CFP walking distance increases from age stage 3 to age stage 5. The growth rate from age stage 4 to age stage 5 is relatively large. The travel distance decreased overall from age stage 6 through 8 (see Fig. 7).

5.4 Changes in Travel Mode with Age

Public transit is the first choice for pilgrim travel, with some differences between age groups. From age stage 3 to age stage 6, the proportion using bus travel reached 63%. This high value is supported by the policy that people aged 60 and above can use unlimited free public transportation. Due to the difficulty in getting on and off, people in age stages 7 and 8 tend to choose the bus less often. There is a download trend in the use of walking from age stage 5. The selection of private cars as a travel mode first rises and then falls; drivers in age stages 3 and 4 convert into riders in age stages 5–8.

6 Analysis of Travel Chain Based on Pilgrimage

This study assumes that pilgrimage behavior occurs within the full travel chain. The travel chain is defined as the travel from the start of the residence, to the return to the end of the trip; it includes the full trip, which may be composed of a number of activities or travel segments. The origin and destination (OD) points of the travel chain are defined by the fixed settlements of the pilgrims. For analysis purposes,

Table 5 Generation of travel chain for pilgrims, under the premise of pilgrimage

Population types	Employed			Not employed		
	Number of travel chains (piece)	Number of travelers (person)	Incidence (piece/person day)	Number of travel chains (piece)	Number of travelers (person)	Incidence (piece/person day)
The pilgrims	1135	1456	0.78	1097	1230	0.89

the pilgrims were divided into two groups: employed and not employed (including retirees). This paper analyzes the travel chain characteristics for both these groups; the results help analyze relationships between travel behavior, economic status, and living conditions.

6.1 Travel Chains of the Pilgrims

The daily average travel chain incidence for pilgrims is 0.83 segments/person day, and the daily home-based travel chain incidence of pilgrims who are not employed is 0.89 segments/person day. Due to work constraints, the pilgrimage-based travel chain incidence for employed pilgrims is 0.78. Pilgrims who are not employed have no work restrictions, and therefore, have more freedom with travel arrangements. For pilgrims whose primary travel purpose is pilgrimage, the travel chain incidence is more than 0.8. This means that pilgrimage has been integrated into the pilgrims' life. Table 5 shows the specific conditions associated with travel chains.

6.2 Gross Travel Time Distribution of Travel Chain of the Pilgrims

There are significant differences in the gross travel time in the travel chain of employed and not-employed pilgrims. Employed pilgrims spend more time at work, and spend more gross travel time in the travel chain than unremployed pilgrims. Employed pilgrims tend to spend 5–6 h Per day while the unemployed pilgrims trend to spend 4–5 h, as shown in Fig. 8.

Lhasa is located in the plateau region, many people may tire after 5 h of continuous activity, which explains why this number of hours appears to be a limit. In addition to work and the pilgrimage, pilgrims also need some social activities. More employed pilgrims have activities that last 7–8 h than not employed pilgrims. This suggests that the physical index and economic condition index of employed pilgrims are higher than those of not-employed pilgrims. The majority of not employed pilgrims are more than 60 years old, while most employed pilgrims are under the age of 60.

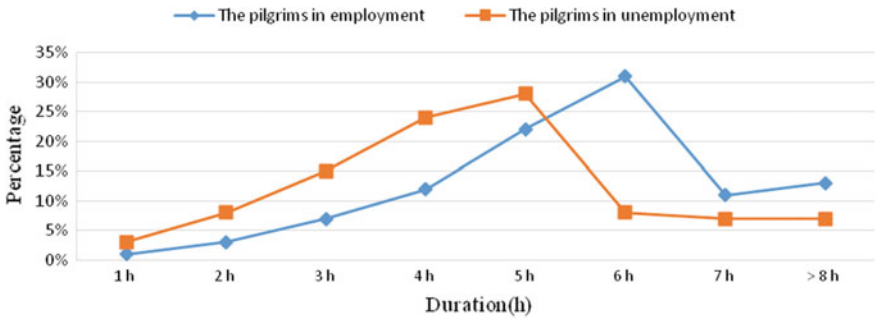


Fig. 8 Chart showing the distribution of the pilgrims' travel chains

6.3 Analyzing the Purpose(s) of the Pilgrims' Travel Chain

The purpose chain points to a travel chain that includes pilgrimage; the travel purpose drives the travel chain.

The formula of a pilgrim's travel chain:

$$\Phi = H + \alpha \sum_{i=1}^{n} P_i + pi + \alpha \sum_{i=1}^{n} P_i + H \tag{2}$$

H: Positions the family as the target, or the arrival of; Φ: The collection of travel chain purposes; pi: Travel purpose for joining the pilgrimage; p_i: The purpose of the entire trip chain, in addition to the pilgrimage, work outside N0. i purposes, $i = 1, 2, 3 \ldots, n$; α: The parameters of the travel destination, the value is fixed at 0, 1.

The travel purposes of the general travel chain are the focus of research. S represents the purpose of shopping; medical treatment is represented by H; recreation and entertainment is represented by L; living activities such as to buy food are represented by Li; and other purposes are captured using O.

The travel chain of pilgrims is different from other populations; pilgrimage must be a purpose in the process. Figure 9 shows the differences in travel chain composition between employed and not employed pilgrims.

As Fig. 9 shows, a diverse life includes diverse travel purposes, so employed pilgrims show more purposes in the travel chain than those who are not employed. Not employed pilgrims had a higher proportion of living activities, while shopping and leisure entertainment accounted for a larger proportion of travel for employed pilgrims.

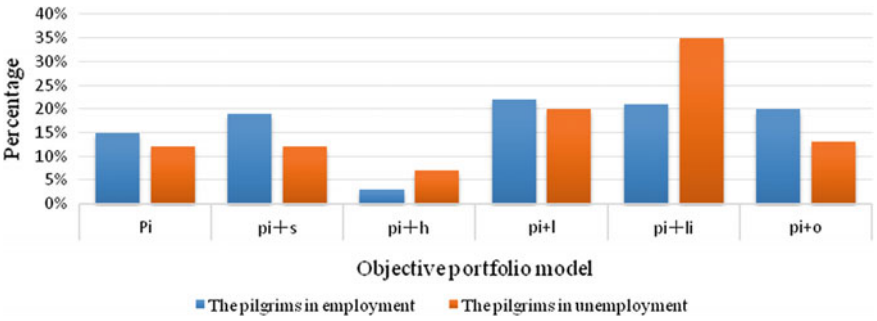


Fig. 9 Chart of distribution of pilgrims' travel chain

7 Conclusions

Based on the intermittent observations from 2011 to 2016, and questionnaires and interviews in 2016 and 2017, this paper draws the following conclusions.

(1) Pilgrims in Lhasa have an average travel rate of 2.25, which is close to the 2.22 rate of non-pilgrims. In daily life, Lhasa pilgrims mainly take public transport (67%) and walking trips (11.7%). However, during religious circumambulation, walking is the main choice for travel. The travel time for both pilgrims and non-pilgrims generally falls between 21–40 min, with departure times generally falling between 6:00–12:00. The morning rush hour for pilgrim travel is approximately one hour earlier than non-pilgrims.

(2) Residents between the ages of 24–48 represent the most active travel group, with a travel rate of 2.43 per day. Residents over 70 have an active travel rate of only 1.83, demonstrating that the travel rate for pilgrims over the age of 48 decreases with age. The amount of time spent on travel increases with age, with people from 61 to 65 covering longer distance than others, at a maximum of 7.8 km. After that age, travel time begins to decline to 3.9 km for people over 75. People aged 24 to 48 travel the shortest distances. With respect to travel mode, the bus is the first choice for people of all ages; however, it is more popular among younger people than older people. The opposite is true for private car use. Policies promoting free bus rides for people over 60 in Lhasa have promoted bus use; however, the bus is not comfortable, so people over 70 tend to use other vehicles, such as private cars. This has made the private car a popular vehicle among pilgrims. Travel mode trends reflect the natural aging of the population. Research comparisons between the two age groups (ages 24–60 and 61–75) indicate that older people are more active in religious circumambulate.

(3) With an average travel chain rate of 0.83 per day, the rate for the employed group is 0.78 per day, which is far below the not-employed group, at 0.89 per day. Religious travel accounts for a large proportion of travel in the daily life of pilgrims, suggesting that religious activities are an indispensable part for life. More employed people experience a total travel time of more than 6 h

than not-employed people. People travelling a total of 5 h account for more than 20% of people across all time periods; the distribution between employed and unemployed people is 22% and 28%, respectively. With respect to travel purpose, the not-employed people spend more travel time on life activities; employed people have more diversified travel purposes.

(4) As a special group, pilgrims have different travel patterns than other groups. Some factors, such as climate, age, and religion, impact pilgrim travel; employment has a minimal impact. During their travel, pilgrims follow both laws of nature and religious rules.

(5) Pilgrimage is an essential part of daily life; as such, the purpose of travel has become increasingly diverse. Pilgrim travel benefits significantly from the development of public transportation; this is demonstrated in the travel rate index, travel time, travel chain rate, and duration of the travel chain. Reasonable planning and management of urban roads and circumambulation roads can help serve the daily needs of pilgrims.

Acknowledgements We thank the Center for Tibetan Studies of Tibet University and the Bus Company of Lhasa Public Traffic Group for valuable data. The work was supported by the Humanities and Social Sciences Foundation of the Ministry of Education in China (Grant No. 17YJAZH010).

References

1. Zhang H (2013) The characteristics of group behavior of pilgrimage practice of circumambulation in Tibetan Plateau. J Tibet Univ (Social Science) 1:196–201
2. Karmay SG (1998) The arrow and the spindle: studies in history, myths, rituals and beliefs in Tibet. Mandala Book Point
3. Zheng Q (2008) Cultural anthropology of pilgrimage and tourism. J Soc Sci Hunan Normal Univ 4:96–99
4. Tian Q (2016) Analysis of ecological pilgrimage of Tibetan Buddhism. J Yunnan Inst Socialism 76–83
5. Fu J (2016) Religious leisure and identity construction of Tibetan and Non-Tibetan Pilgrims—a case study of Tibetan Mountain Pilgrimage. Zhejiang University
6. Zhang H, Chen Y (2012) On a Tibetan Religious practice "circumambulation" and personal religious experience. J Univ Tibet (Social sciences Edition) 1:125–130
7. Peng W, Shiming Chen (2014) Tibetan pilgrimage in the process of social change: a case study of Jiuzhaigou. Qinghai Ethnic Stud 3:1–7
8. Zhang Z (2009) Travel behavior characteristics with analysis for the elderly. Beijing Jiaotong University
9. Zhang M et al (2008) Travel behavior analysis of the females in Beijing. J Transp Syst Eng Inf Technol 8(2):19–26
10. Qu D, Zhuang J, Liu T, Nakatomi I, Yang M, Wu Y (2001) Analysis on the resident trip characteristics and study on the transport development policies in Suzhou. J Southeast Univ (Natural Science Edition) 31(03):118–123
11. Ye M, Yu M, Guo X, Dou X (2014) Analysis of effects of contributing factors on choice of activity pattern in historic urban areas. J Southeast Univ (Natural Science Edition) 44(1):211–215

12. Srinivasan S, Bhat CR (2008) An exploratory analysis of joint-activity participation characteristics using the American time use survey. Transportation 35(3):301–327
13. Kang HJ, Scott DM (2011) Impact of different criteria or identifying intra-household interactions: case study of household time allocation. Transportation 38(1):81–99
14. Mosa AI, Esawey ME (2013) An investigation of household interactions in daily in-home and out-of-home maintenance activity participation and social behavior in Cairo, Egypt. Transp Lett Int J Transp Res 5(4):201–212
15. Wheatley D (2014) Travel-to-work and subjective well-being: a study of UK dual career households. J Transp Geogr 39:187–196
16. Hsu HP, Saphores JD (2014) Impacts of parental gender and attitudes on children's school travel mode and parental chauffeuring behavior: results for California based on the 2009 National Household Travel Survey. Transportation 41(3):543–565
17. Bueno PC, Gomez J, Peters JR et al (2017) Understanding the effects of transit benefits on employees' travel behavior: evidence from the New York-New Jersey region. Transp Res Part A Policy Pract 99:1–13
18. Cheng G (2016) Research on the public transport organization during Shoton Festival in Lhasa. Highway 1:162–166
19. Wang Z, Cheng G (2014) Analysis on the optimization of public traffic services during the religious festivals in Lhasa. J Univ Tibet (Natural Science Edition) 2:95–101
20. Chen G (2006) An analysis of the holy land complex of Tibetan Buddhist Pilgrims. Religious Stud (01):182–186

Research on Drivers' Cognitive Level at Different Self-explaining Intersections

Wuhong Wang, Shanyi Hou, Xiaobei Jiang and Qian Cheng

Abstract One demand for road is the ensurance of self-explaining, under which means road users can make correct subjective classifications and expectations of road environment. Based on quantification of driver's driving cognitive behavior and the self- explaining road theory, this paper designs road environments with different self-interpretation levels as experimental scenes. Through a driving simulation experiment, the changing process of driver's cognitive workload level is simulated based on Hidden Markov Model. The Hidden Markov Model identifies the driving intention under the combined working conditions, thereby judging driving awareness of the road environment, and evaluating the self-interpretation level of each experimental scene.

Keywords Urban road intersection · Drivers' cognitive level · Self-interpretation feature · Double HMM

1 Introduction

After analyzing about driving reliability, Indiana University [1] found that the driver had the most traffic accidents caused by cognitive errors during driving. Cognition, as a source of all external information, plays a vital role and is the core process of driving behavior. At present, many scholars have studied the impact of different road environmental parameters on driving behavior. Mazet et al. [2] pointed out that drivers will subjectively classify roads, and wrong road classification will cause incorrect driving expectations. Causes the driver to make inappropriate driving behavior. The driver's perception of the traffic environment is due to the driver's classification of the traffic environment is not based on the individual environment, but an abstract feature of a set of basic attributes that can cause the driver's driving expectations. Many traffic accidents are caused by drivers' false expectations of the road environment. Road

W. Wang (✉) · S. Hou · X. Jiang · Q. Cheng
School of Mechanical and Vehicular Engineering, Beijing Institute of Technology, Beijing 100081, China
e-mail: wangwh@bit.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_64

traffic safety research needs to change from the "driver-traffic accident" model to the "road design-driver-traffic accident" model urgently.

2 Self-explaining Road

The concept of self-explaining roads originally originated in the Netherlands, Germany, Denmark and other countries in Europe. In 1995, Theeuwes [3] proposed the concept of Self-explaining Road (SER): a road system that only uses road design to encourage road users to make safe driving behavior. Self-explaining roads can help drivers establish correct subjective road classifications by changing the visual characteristics of the road, thereby affecting the driver's driving behavior, reducing driving mistakes, and improving safety.

2.1 Self-explaining Intersection Experiment

The self-explaining intersection is an environment that is based on the drivers' driving cognitive characteristics during the design process, guiding the driver to select the appropriate driving behavior. The self-explaining intersection design can change the driver's cognition degree through road design, such as lane width, road marking, traffic signs and other road features, guiding the driver's perception and road environment perception, guiding driving behavior directly or indirectly [4].

In this experiment, the experimental scene will be built by UC-win/Road software, and the driver's movement and eye tracker will be used to collect the eye movement data such as vehicle speed, trajectory map, gaze point and gaze area during driving [5].

In this experiment, 50 experimenters of different ages and driving ages were randomly selected from Beijing Institute of Technology as experimental samples. The average age of the subjects was 29.1 years old and the standard deviation was 10.52. The selected drivers all held driving qualification certificates and had driving experience.

2.2 Experimental Scene

In this experiment, five different self-explaining traffic scenarios were designed, and in order to avoid the cognitive load caused by the self-interpretation levels change between different scenes, the driver was overloaded, and a transitional section was designed between adjacent scenes to slow down the driving cognitive change. This experiment uses UC-win/Road software to create the required traffic scenarios.

The five self-explanation levels of traffic scenes are different, and the purpose is to study the drivers' cognition by adding or improving different road environment parameters. The five traffic scenes include the common intersection environment, road environment, the intersection environment and road environment with self-explaining characteristics. The two groups form a comparison group to study the application effects of different self-explaining traffic scenarios further.

1st intersection	2nd intersection	3rd intersection	4th intersection	5th intersection
Two-way two-lane deceleration intersection	The entrance road sets the community island to slow down	Ordinary two-way four-lane signal control intersection	Two-way four-lane signal control intersection with optical illusion deceleration marking	Two-way four-lane signal control intersection with diversion lines and changing the color of the bicycle lane

3 Driving Intention Identification Based on Double Hidden Markov Model

3.1 Experimental Data Processing

In this experiment, normal steering, emergency steering, and straight-line driving were selected as the inspection conditions, and the sensor data collected by the experiment was constructed into the data set of the entire HMM model. These data are processed according to the following process.

(1) Firstly, the experimental data is filtered and amplified, and the processed data is divided according to its use: vehicle speed, pedal information, and steering wheel information. Subsequently, the classified information is divided and classified so that each data segment corresponds to short-term driving behavior data. Then the t-detection method is used to combine the characteristic parameters of different data segments to eliminate the abnormal data.
(2) Secondly, the limit value between the normal operation and the emergency operation in the driver's driving consciousness is determined by the K-means algorithm to help verify the correctness of the recognition result.

In this experiment, the threshold value of emergency steering, normal steering, and driving operation during straight-line driving is determined by the Matlab program based on the K-means algorithm. The result is shown in Fig. 1.

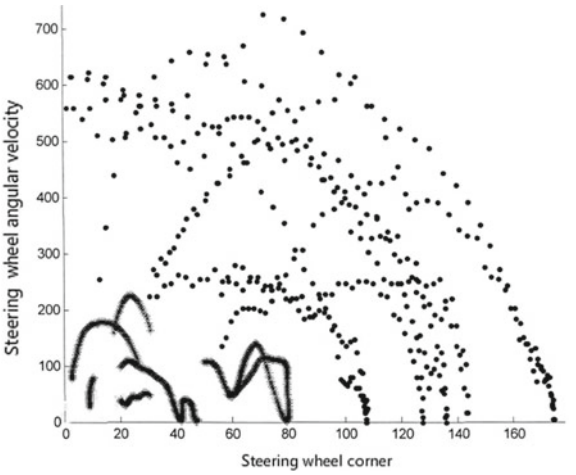


Fig. 1 Normal steering, emergency steering, and limits for driving operations during straight-line driving

3.2 Offline Training

Determining model parameters is an important step in a two-layer HMM model. In this experiment, all parameters in the double-layer HMM model can be trained offline according to the driving operation signal. The multi-dimensional Gaussian HMM model and multi-dimensional discrete HMM model after training are loaded into matlab, and LabVIEW is used for online identification verification.

Taking the multi-dimensional discrete HMM model of normal turning as an example, the parameters obtained are as follows:

$$\pi 1 = \begin{bmatrix} 0.918 \\ 0.017 \\ 0.004 \end{bmatrix} \quad \mathrm{A1} = \begin{bmatrix} 0.9877 & 0.0231 & 0.0002 \\ 0.9173 & 0.0691 & 0.0014 \\ 0.9352 & 0.0426 & 0.0212 \end{bmatrix}$$

$$B1 = \begin{bmatrix} 0.1507 \cdots 0.4008 \\ 0.0842 \cdots 0.2342 \\ 0.0003 \cdots 0.0021 \end{bmatrix}_{3\times 8} \quad \mathrm{B2} = \begin{bmatrix} 0.0399 & 0.8403 & 0.0103 \\ 0.0067 & 0.8919 & 0.0016 \\ 0.8649 & 0.2336 & 0.0016 \end{bmatrix}$$

$$\mathrm{B3} = \begin{bmatrix} 0.0485 & 0.9515 \\ 0.0552 & 0.9648 \\ 0.0007 & 0.9993 \end{bmatrix}$$

3.3 Lab VIEW Online Verification

The LabVIEW program is programmed to drive the data stream. In the main loop, the program collects six sensor data in the driving simulator through the dynamic link library, and amplifies and filters the collected data, and then processes the queue technology to transmit the processed data to From the loop. After receiving the data from the loop, it is sorted and sent to the corresponding module (pedal module, steering module and speed classification module). After being identified, a three-dimensional identification result string is obtained and sent to each intent layer HMM model to display the identified driving intention result. As shown in Fig. 2, it is the block diagram built in the LabVIEW program.

In this experiment, a large number of online identification experiments set the limit values for three kinds of compound working conditions: normal turning, emergency turning, and straight-line driving. When the limit value is exceeded, this condition occurs. For the selection of time steps from the loop, 0.08 s is usually selected as the execution time step to obtain higher accuracy. The multi-dimensional Gaussian HMM model of this experiment has done a total of 300 pedal recognition, 100 steering wheel steering identification, the identification accuracy can reach 99.5%.

As shown in Table 1, this table is the identification result of the steering module considering only the operation layer HMM model. The correct identification result is displayed on the yellow bottom.

A set of data is randomly selected from the collected experimental data for driving intention identification, and the results are shown in Figs. 3 and 4 and Table 2.

For this section of driving operation data, the period of the latter two driving intent identification models indicates that the driver is in the accelerated braking phase, that is, in the emergency turning phase.

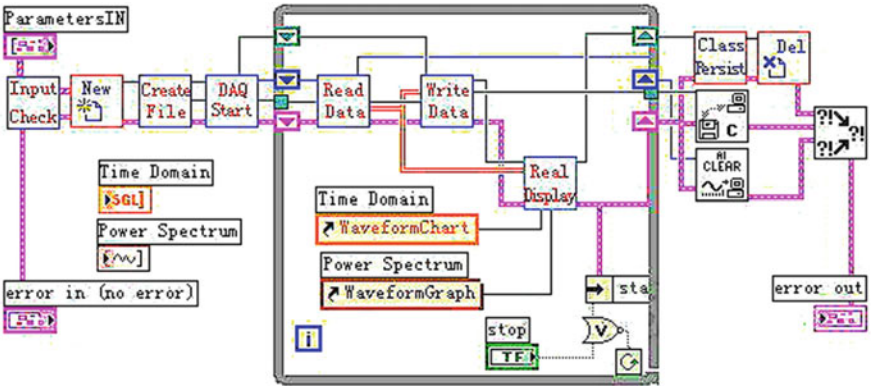


Fig. 2 Lab VIEW block diagram of the driving intent identification model

Table 1 Operation layer steering module identification results

	Normal steering HMM model	Emergency steering HMM model	Emergency steering HMM model
Normal steering 1	−35,065	−7426	**−1344**
Normal steering 2	−24,107	−3020	**−949**
Normal steering 3	−10,775	−6372	**−1400**
Normal steering 4	−18,719	−8335	**−4521**
Emergency steering 1	−21,865	**−1249**	−5251
Emergency steering 2	−28,052	**−1859**	−17,912
Emergency steering 3	−74,351	**−1790**	−38,526
Emergency steering 4	−68,507	**−1820**	−31,266
Straight line 1	**−929**	−6462	−21,873
Straight line 2	**−1801**	−5491	−8392
Straight line 3	**−1686**	−14,763	−35,687
Straight line 4	**−564**	−7770	−17,735

Bold represents the set of data is randomly selected from the collected experimental data for driving intention identification

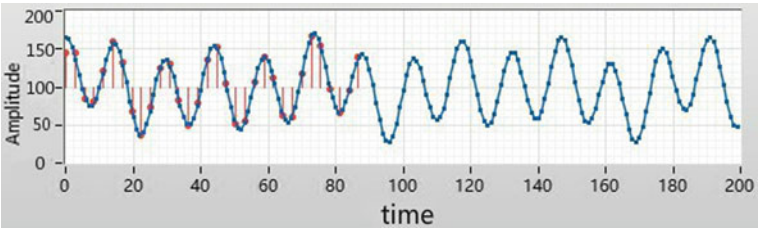
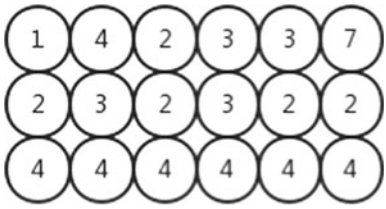


Fig. 3 Steering wheel operation steering data identification result

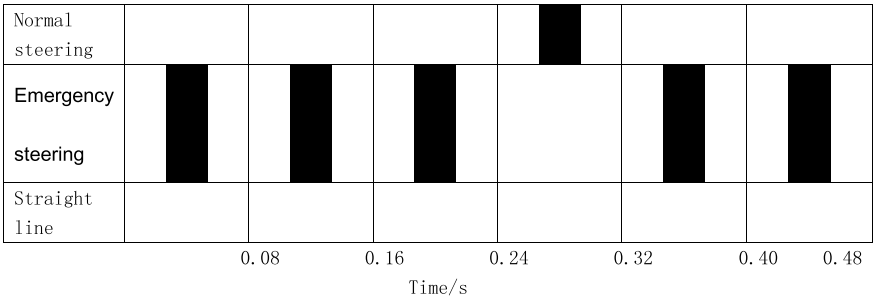
Fig. 4 Identification result of consciousness layer HMM



3.4 In Conclusion

5 experiments were carried out on the intentional layer HMM model under normal compounding, emergency turning and straight-line driving under three composite conditions. The accuracy of the obtained identification results shows that the accuracy of the two-layer hidden Markov model is obtained. As well as real-time, it can well

Table 2 Identification result of operation layer HMM



identify and predict the driver's driving intentions. The driver's cognition is judged by the identification of the driving intention.

In time and space, drivers first perceive the road environment and generate a corresponding driving intention—driving awareness. In this paper, a two-layer hidden Markov model is constructed. Through the analysis and analysis of the driving behavior characteristic information and the intrinsic connection between driving intention and driving behavior, the driving intention recognition of normal steering, emergency steering and straight driving is realized. The model can verify the driver's perception of the road environment and determine the driver's driving awareness level of his or her road environment.

It can be seen from Figs. 5, 6, 7, 8 and 9 that the driver's trajectory is chaotic at the intersection 1, 3, and 4, and the trajectory of the vehicle crossing the 2, 5 is obviously better. It is proved that the self-interpretation design of the roundabout and the diversion line obviously has a certain degree of influence on the driver's cognitive level, which can help the driver to better perceive the curve and restrain the driving trajectory to a certain extent.,the correctness of the double-layer hidden Markov model for driving cognitive level is further proved.

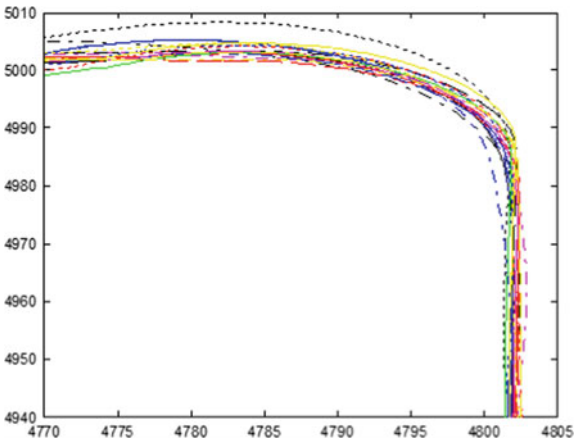


Fig. 5 The date of driver's vehicle track

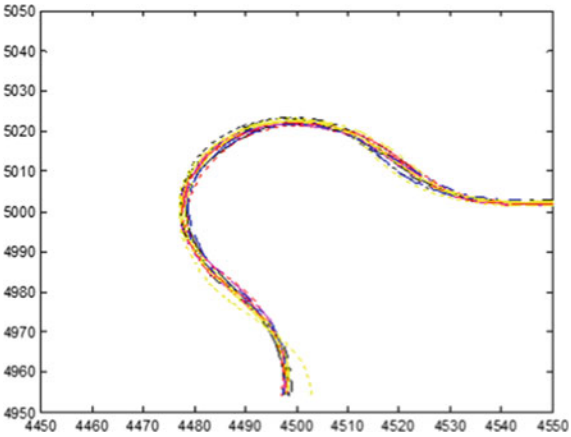


Fig. 6 The date of driver's vehicle track

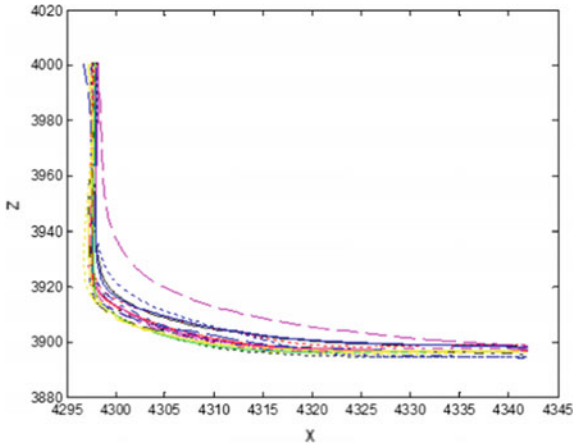


Fig. 7 The date of driver's vehicle track

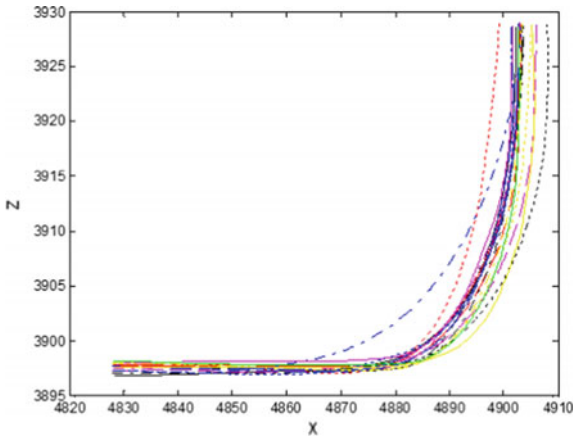


Fig. 8 The date of driver's vehicle track

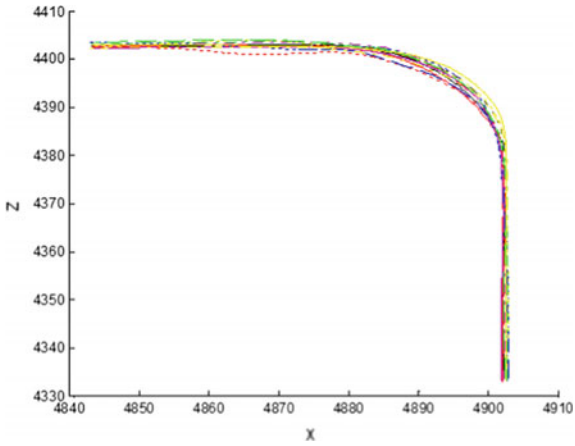


Fig. 9 The date of driver's vehicle track

4 Conclusions

(1) Introduce the concept of self-explaining road, and put forward the concept of self-explaining plane intersection based on driving cognitive characteristics and self-explaining characteristics. The related concepts of self-explaining plane intersections are expounded, and self-interpreting plane intersections are proposed. Fault-tolerant design theory and design methods.

(2) Based on the characteristic data of different data sets such as pedal speed and steering wheel angle under different composite conditions, the driving intention identification model of double-layer HMM model is constructed and the model is verified online. The purpose is to verify the driver's perception of the road environment by identifying the driver's intention and to determine the driver's driving awareness level of his or her road environment.

References

1. Nagayama Y (1978) Role of visual perception in driving. Iatss Research
2. Mazet C, Dubois D (1988) Mental organizations of road situations: theory of cognitive categorization and methodological consequences. In: SWOV conference: traffic safety theory and research methods. Amsterdam (1988)
3. Theeuwes J, Godthelp H (1995) Self-explaining roads. Saf Sci 19(2–3):217–225
4. Brackstone MA, Waterson BJ (2004) Are we looking where we are going? An exploratory examination of eye movement in high speed driving. In: Meeting of the Transportation Research Board
5. Ranney TA (1994) Models of driving behavior: a review of their evolution. Accid Anal Prev 26(6):733

A Novel Multiple Object Tracking Algorithm for Autonomous Vehicles

Hai Deng, Ming Gao, Li-sheng Jin and Bai-cang Guo

Abstract Multiple object tracking is a vital task for autonomous vehicle environment perception. In this paper, we design a novel multi-object tracking method for autonomous vehicles. In the detection section, we utilize popular Faster-RCNN as our baseline method. Then, in data association, we combine appearance, motion, and interaction model to build a unified feature descriptor to explore the nature of tracking object. We evaluate our algorithm on a popular and standard benchmark and compare with the state-of-the-art methods. The results denote that our algorithm achieve good performance at high frame rates.

Keywords Autonomous vehicles · Computer vision · Multiple object tracking · Data association · Convolutional neural network

1 Introduction

Environment perception, aiming to recognize and understand the surroundings, is one of the most fundamental tasks of autonomous vehicles [1]. The content of environment perception can be extended as many subtasks, such as semantic segmentation [2], object detection [3], object tracking [4] and scene reconstruction [5]. Although many researchers have devoted to object detection, the real-time and high-efficacy detection problem is still challenging due to the weaker compute power of vehicle

H. Deng (✉)
Integration Platforms Department, National Engineering Research Center of Railway Vehicles, Changchun 130022, China
e-mail: denghai@cccar.com.cn

M. Gao · L. Jin · B. Guo
Transportation College, Jilin University, Changchun 130022, China
e-mail: 794042627@qq.com

L. Jin
e-mail: jinls@jlu.edu.cn

B. Guo
e-mail: m18646323411@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_65

computing unit and complex situations. Thus the object tracking can be a good trade-off solution to consummate the object detection methods. Moreover, object tracking also can be applied to behavior analysis [6], motion estimation and another high-level computer vision task. In this paper, we pay attention to multi-object tracking (MOT), which can obtain all predefined object location in a complex situation. Especially, we design our algorithm for difficult and time varying traffic scenes, which heavy occlusion, background clutter, fast motion, deformation occur frequently that may hamper the tracking process.

Most of popular MOT trackers follow the tracking by detection pipeline that divide the tracking process as two parts: object detection and data association. In the field of object detection, mellow and approved object detectors, such as DPM [7], Faster-RCNN [8], YOLO [9] and SSD [10], perform better accuracy to locate object at the pixel-level to feed to the data association. However, the detectors may introduce false alarms, miss detections and low overlap detections, which both make the problem much more difficult. Then, data association period focus on connecting the same object at different frames to construct some tracklets and final trajectory. Many data association methods are modeled as graph-based combinatorial optimization algorithms that the tracking objects are nodes and the connected paths are edges. Because of the above optimization algorithm can be formulate as many kinds of problems, such as minimum cost network flow, conditional random field (CRF), graph labeling and dynamic programming, are usually NP-hard problem.

For better understand above data associations, we introduce some typical data association methods. Zhang [11] introduce a popular global information based min-cost flow data association method, which integrated explicit occlusion model (EOM) to deal with occlusion constraints. Then, many network flow methods were proposed, Pirsiavash [12] proved that the global solution of network flow can be obtained by a greedy short path computation. By modeling the MOT tracker as bipartite graph matching, many interesting algorithms were presented. Breitenstein [13] proposed a particle filtering framework to robust track on fierce multi-person moving under complex scenes. Shu [14] reckoned the part-based model and person-specific SVM will capture the real-time changing appearance of human body. The dynamic programming methods also perform the state-of-art performance on MOT challenge. Due to the wide application of dynamic programming, almost all the basic dynamic programming methods can be utilize as the data association module, such as, linear programming [15], quadratic programming [16] and K-shortest paths [17]. Note that subgraph multicut [18] methods become the new fashion before the deep learning methods, which can link and cluster plausible detections across space and time as a subgraph decomposition problem. Another graph-based data association method is conditional random field (CRF). The energy minimization methods [19] define a graph where the detections are nodes and the tracklets are edges, the labeling problem is discrete-continuous which include many terms to model the pedestrian tracking process.

In this paper, we propose a novel multiple object tracking algorithm for autonomous vehicles under real and complex traffic scenery, in particular for pedestrian tracking. With the development of deep learning, convolutional neural networks

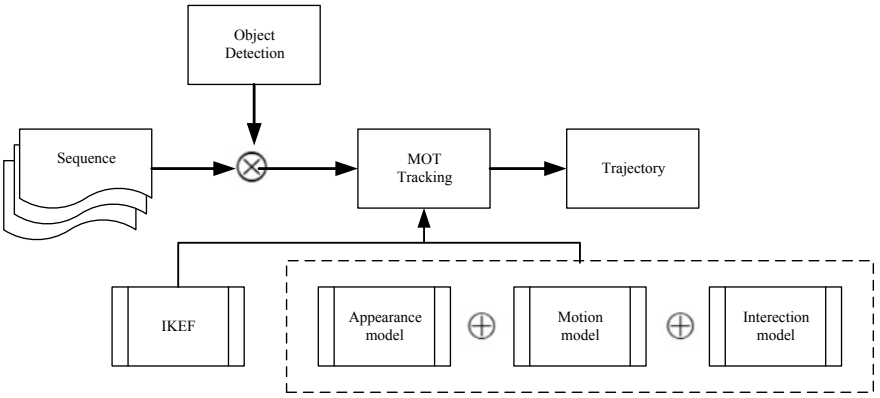


Fig. 1 The pipeline of proposed MOT tracker

seem to be the standard necessity of computer vision tasks, like object classification, video object recognition, just mentioned few. With this paradigm, we also make good use of convolutional neural networks to present a deep learning framework for online multiple object tracking which can operate more than 25 FPS on a single CPU. The main intuition of our algorithm is that single CNN appearance descriptor cannot cope with complex situation. We utilize CNN appearance with motion model to make good use of convolutional features and motion features. Deploying and calculating different pre-train model can explore the nature of long and short dependency of appearance model. Another merit of our algorithm can be summarized as the low-complexity data association. Traditional data association method is always modeled as combinatorial optimization problem, which seldom can find the global solution in polynomial time. In order to come up with this issue, we modified the classical Kalman filter and with a Hungarian algorithm. The pipeline of this algorithm can be seen in Fig. 1.

The rest of this paper is organized as follows: Section "Methodology" presents the proposed main methods used for multiple object tracking in crowded and complex scenes. Section "Experimental Results" shows the experiments' results based on the popular MOT benchmarks. At last, Section "Conclusion" summarizes the study and discusses future works.

2 Methodology

The MOT tracker component can be divided as two parts: object detection and data association. There are many excellent object detectors for decades. Deformable Parts Model (DPM), which can be seen as extend of Histograms of Oriented Gradients (HOG), is one of the best object detection algorithms before deep learning era. To benefit from the advancement of CNN based detection, we adapt the Faster Region

CNN (Faster-RCNN) as our baseline detection framework. In every single frame, the detection algorithm operate two stages, which the first stage can get region of interest (ROI) for the next stage to estimate and locate the object. Then the detection results feed to our data association section. For a better understand of this detection architecture, we recommend the readers to read relational papers [8].

State Estimation

Before introduce our state estimation algorithm, we assume that our tracking algorithm is deployed on a monocular camera and we have no priori information. We utilized the Invariant Extend Kalman Filter (IEKF) [20] to track object. The state space s of our algorithm denote as eight dimensional $s = (x, y, w, h, x_v, y_v, r, t)$, where (x, y, w, h) is the bounding box location, (x, y) is the left-top pixel coordinate and (w, h) is the width and height of object, (x_v, y_v) is the component velocity, r is the scale change, t is the time stamp, which both contain the static and dynamic variable. The IEKF is suitable for our task, which the velocity motion is not constant and the observation model is non-linear, make the. The original Kalman Filter (KF) can be written as follow:

$$s_k = f(p_{k-1}) + \mathrm{w}_{k-1} \tag{1}$$

$$z_k = g(p_k) + v_k \tag{2}$$

$s_k \in \mathrm{R}^n$ is the state vector at time k, $z_k \in \mathrm{R}^n$ is the observe vector. The non-linear function $f(\cdot)$ and $g(\cdot)$ are the dynamic model function and observe model function. w_k and v_k are process noise and observe noise respectively, which is irrelevant.

Assignment Problem

We model the data association as an assignment problem, which use the Hungarian algorithm for our IKEF state estimation. The time complexity of the algorithm is $O(n^3)$, the Hungarian algorithm can be seen as a fast assignment algorithm. The predicted variable and the new coming observe value can be design as numerical matrix to assignment the object of different frames. We learn and formulate the motion, appearance, and interaction cue through fuse these three types of information.

Motion information aims to capture the inner nature of interested objects, mainly pedestrians. In order to incorporate motion information, we utilize the squared Mahalanobis metric to calculate the similarity between Kalman states and newly coming measurements:

$$d^{(1)}(m, n) = (d_m - d_n)^T S_i^{-1} (d_j - y_i) \tag{3}$$

where we denote the projection of the m-th tracking into measure space by (Y_m, S_m) and the n-th bounding box result by d_n. Benefiting greatly from the Mahalanobis distance, we can know the state estimation uncertainty by measuring the number of standard deviations detection away from the mean track location. Moreover, the

advantage of this metric can eliminate the unlikely associations by defining the Mahalanobis distance at different confidence interval, which can calculate by χ^2 distribution. In this paper, we set the confidence interval as 95%, which is common used. The indicator as follow:

$$b_{m,n}^{(1)} = 1[d^{(1)}(m,n) \leq t^{(1)}] \tag{4}$$

when the evaluate is 1 denote the association between the m-th and n-th is trustable.

While the Mahalanobis distance and IEKF can present the motion information of objects, we still need the appearance to make long-time decency. Traditional appearance model includes raw pixels, HOG, Color Name, and so on. However, these features cannot save the image spatially information, which the convolutional feature perform well on spatially operation. For every object detection d in picture, we calculate the appearance descriptor A. We set the max number of objects in a frame is 50. And the metric measures the Jaccard distance between the m-th and n-th detection in appearance domain:

$$d^{(2)}(m,n) = 1 - \frac{|A_m \cup A_n| - |A_m \cap A_n|}{|A_m \cup A_n|} \tag{5}$$

Then, we also design a binomial variable to perform the matching degree:

$$b_{m,n}^{(2)} = 1[d^{(2)}(m,n) \leq t^{(2)}] \tag{6}$$

In our experiment period, we apply transfer learning to fine-tune a pre-trained detection network, which can compute the appearance descriptors. We will describe the detail in the next experimental section.

The motivation of our interaction model derives from the observation that in crowded pedestrian situation, for example, the pedestrians rush to the opposite side on the sidewalk. The pedestrians do not like to close each other too much, even in very crowded situation. We incorporate an Explicit Occlusion Model (EOM) on the tracking objects. The EOM can engender occlusion hypotheses with sort of occlusion constraints.

EOM include two parts: *hypothesis generate* and *hypothesis apply*. The hypothesis generate section assure that only if the distance $|x_m - x_n|$ and the scale variation $\frac{s_m}{s_n}$ are in a certain interval, written as:

$$O_{m,n} = (|x_m - x_n|, \frac{s_m}{s_n}) \tag{7}$$

$O_{m,n}$ denote as the set of objects m and n, at every time the value of $O_{m,n}$ beyond a certain threshold to generate a hypothesis.

Then, the maximum-a-posteriori (MAP) apply to evaluate the set $O_{m,n}$. The occlusion constraints are introduced as:

$$f_m^n \leq F(t), \forall x \in O_{m,n} \tag{8}$$

where f_m^n denote the occlusion function and $F(t)$ is the occlusion threshold function, f can only be 0 or 1, these constraints make the rest hypotheses are true.

Finally, we combine the three parts of our model to a unified assignment problem, as follow:

$$c_{m,n} = \lambda_1 d^{(1)}(m,n) + \lambda_2 d^{(2)}(m,n) + (1-\lambda_1-\lambda_2) f_m^n \tag{9}$$

λ_1 and λ_2 are both weight, which range from [0,1]. The larger the number is, the more important of this attribute is.

3 Experimental Results

Our algorithm created and operated on a desktop PC. The code was implement on Windows Tensorflow written in Python and we accelerate our algorithm on an Inter (R) Core (TM) i7-4790, CPU @ 3.60 GHz, 16 GB RAM, and Nvidia GTX 1080ti GPU. Based on many tentative experiments, the weight of feature parameters is setted as $\lambda_1 = 0.4$ and $\lambda_2 = 0.4$. And the deep appearance descriptor we utilized is a lightweight and modified YOLO [9] detection network. The network architecture has 8 convolutional layers, including common convolutional layer, max-pooling layer and fully-connected layer. We omit the pad in every layer, and the basic network framework shown in Table 1.

Multiple Object Tracking (MOT) Benchmark is one of the most popular video multi-object tracking datasets. With this benchmark we can understand our method in a more meaningful quantification unified framework. We evaluate our algorithm on MOT16 benchmark. This benchmark includes 14 challenging video sequences, 7 training sets and 7 test sets, both contain outdoor and indoor sequence with moving

Table 1 The network architect of our appearance descriptor

Name	Filter size/stride	Output size
Original image		1920*1080*3
Conv 1	3*3/2	960*540*12
Max pool 1	2*2/2	480*270*48
Conv 2	3*3/2	240*135*96
Max pool 2	3*3/2	120*65*128
Conv 3	3*3/3	60*30*254
Max pool 3	2*2/2	30*15*512
Fully-connected		128
Batch and ℓ_2 normalization		128

and fixed camera. The benchmark videos detections are generated by popular Faster RCNN [8] object detection algorithm.

We test our algorithm based on CLEAR MOT metrics [21]. The metrics group include the following metrics:

(1) Multi-object tracking accuracy (MOTA): a unified and robust summary to calculate the overall tracking accuracy which include the false positives, false negatives and identity switches.
(2) Multi-object tracking precision (MOTP): the overall tracking precision by calculating the percentage of bounding box overlap between algorithm generate location and ground-truth.
(3) Mostly tracked (MT): the percentage of true value tracks with the same assignment larger than 80% in an object life circle.
(4) Identity switches (ID): the number of algorithms generate detection compare with ground-truth change with each other.
(5) Mostly lost (ML): the frames number that ground-truth tracks under 20% of the life span
(6) Fragmentation (FM): the frames number that interrupted by missing detections.

The results are shown in Table 2. We compare our algorithm with the state of art tracking methods [22–26], both include the batch and online methods. For every index, we use red font to describe the best. Our method achieves state-of-art performance on all the metrics. And the proposed method is better than all the other methods on MOTP, which achieve 81.3. This denote that our algorithm is a robust tracking algorithm. In terms of ML, we also get the value 18.0% with a slightly lower of 1% than the second place. The good results of our algorithm emanate from that our approach can connect the nature of appearance, motion and interaction model.

For better evaluate our algorithm, we also consider speed and qualitative analysis. The overall runtime of above algorithms can be seen in Table 3. From the table, we can find that batch methods seldom can ran faster than 10 FPS. Conversely, almost all the online methods can achieve real-time tracking. During the test stage, our method can run at 35 FPS in a single CPU, which is very suitable for operating on vehicle compute unit.

To evaluate the proposed algorithm on qualitatively performance, we compare this method with other state-of-the-art trackers on different test sets of MOT16. We present the bounding box and ID comparisons in Fig. 2. The images (a) and (b) are from the MOT16-01, which is record people walking around a large square by a static camera. Our algorithm can track people fine. Image (c) and (d) are from the MOT16-03, which denote crowded pedestrian walk at night in elevated viewpoint. The low luminance makes track pedestrians very difficult. However, our algorithm perform slightly slower than common scenes without accuracy loss. Different from other outdoor videos, image (e) and (f) are recorded by forward moving camera in a busy shopping mall. Our algorithm can cope with fast motion in above scene. The image (g) and (h) are extracted by one of the most difficult video sequences MOT16-14. This video is filmed a bus on a busy intersection. The pedestrians in bus

Table 2 Tracking results on the MOT16 challenge

		MOTA↑	MOTP↑	MT↑ (%)	ML↓ (%)	ID↓	FM↓	FP↓	FN↓
KDNT	Batch	68.2	79.4	41.0	19.0	933	1093	11,479	45,605
LMP_p	Batch	71.0	80.2	46.9	21.9	434	587	7880	44,564
NOMT	Batch	62.2	79.6	32.5	31.1	406	642	5119	63,352
MCMOT_HDM	Batch	62.4	78.3	31.5	24.2	1394	1318	9855	57,257
POI	Online	66.1	79.5	34.0	20.8	805	3093	5061	55,914
EAMTT	Online	52.5	78.8	19.0	34.9	910	1321	4407	81,223
SORT	Online	59.3	79.6	25.4	22.7	1423	1835	8698	63,245
Ours	Online	63.1	81.3	31.5	18.0	765	2156	10,982	55,648

Tab.3 Comparison of runtime speed

	KDNT	LMP_p	NOMT	MCMOT_HDM	POI	EAMTT	SORT	Ours
FPS	0.7	0.5	3	35	10	12	60	35

Fig. 2 Qualitatively evaluate on MOT16

stop are crowded and occluded each other. The proposed method can achieve good performance on this video sequence.

4 Conclusion

In this paper, we present a new method for multiple object tracking environment perception. The methods use a popular object detection network to generate detection results. And in data association section, we have employed a triplet feature descriptor, which consist of pre-trained appearance model, linear motion model, and EOM interaction model. Extensive experiments operated on MOT 16 denote that our algorithm can perform better than the state-of-the-art methods and run more than 25 FPS.

Acknowledgements This work is supported by the National Natural Science Foundation (no. 51575229), the National key R & D project of China (no. 2017YFB0102600), and the Electric intelligent vehicle innovation team of the Science and Technology Department of Jilin Province(no. 20180519022JH).

References

1. Zhu H, Yuen KV, Mihaylova L et al (2017) Overview of Environment perception for intelligent vehicles. IEEE Trans Intell Transp Syst 18(10):2584–2601
2. Romera E, Alvarez JM, Bergasa LM et al (2018) ERFNet: efficient residual factorized ConvNet for real-time semantic segmentation. IEEE Trans Intell Transp Syst 19(1):263–272
3. Tomè D, Monti F, Baroffio L et al (2016) Deep convolutional neural networks for pedestrian detection. Sign Proc Image Commun 47(C):482–489
4. Swanson RS, Musa S (1975) The estimation of obstacle and terrain clobber probabilities, AIAA 75-1118. University of Cambridge Press, Cambridge
5. Koltun V, Koltun V, Koltun V et al (2017) Tanks and temples: benchmarking large-scale scene reconstruction. ACM Trans Graph 36(4):78
6. Stratou G, Morency LP (2017) MultiSense—context-aware nonverbal behavior analysis framework: a psychological distress use case. IEEE Trans Affect Comput PP(99):1–1
7. Jin L, Guo B, Jiang Y et al (2018) Study on the impact degrees of several driving behaviors when driving while performing secondary tasks. IEEE Access 6:65772–65782
8. Ren S, He K, Girshick R et al (2017) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149
9. Redmon J, Divvala S, Girshick R et al (2015) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE
10. Liu W, Anguelov D, Erhan D et al (2016) SSD: single shot MultiBox detector. In: European conference on computer vision, p 2
11. Zhang L, Li Y, Nevatia R (2008) Global data association for multi-object tracking using network flows. In: 2008 IEEE conference on computer vision and pattern recognition. IEEE
12. Pirsiavash H, Ramanan D, Fowlkes CC (2011) Globally-optimal greedy algorithms for tracking a variable number of objects. In: IEEE conference on computer vision & pattern recognition. IEEE Computer Society

13. Breitenstein MD, Reichlin F, Leibe B et al (2009) Robust tracking-by-detection using a detector confidence particle filter. In: IEEE international conference on computer vision. IEEE
14. Shu G, Dehghan A, Oreifej O et al (2012) Part-based multiple-person tracking with partial occlusion handling. In: IEEE computer society conference on computer vision and pattern recognition
15. Jiang H, Fels S, Little JJ (2007) A linear programming approach for multiple object tracking. In: 2007 IEEE conference on computer vision and pattern recognition. IEEE Computer Society
16. Dehghan A, Shah M (2017) Binary quadratic programing for online tracking of hundreds of people in extremely crowded scenes. IEEE Trans Pattern Anal Mach Intell 40(3):568–581
17. Choi W, Savarese S (2012) A unified framework for multi-target tracking and collective activity recognition. In: European conference on computer vision. Springer, Berlin, Heidelberg
18. Tang S, Andriluka M, Andres B et al (2017) Multiple people tracking by lifted multicut and person re-identification. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). IEEE Computer Society
19. Andriyenko A, Schindler K (2011) Multi-target tracking by continuous energy minimization. In: The 24th IEEE conference on computer vision and pattern recognition, CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011. IEEE
20. Barrau A, Bonnabel S (2014) The invariant extended Kalman filter as a stable observer. IEEE Trans Autom Control PP(99):1–1
21. Yu F, Li W, Li Q et al (2016) Multiple object tracking with high performance detection and appearance feature. In: European conference on computer vision (ECCV). Springer, Cham, pp 36–42
22. Keuper M, Tang S, Zhongjie Y et al (2016) A multi-cut formulation for joint segmentation and tracking of multiple objects. arXiv preprint arXiv,1607.06317
23. Lee B, Erdenee E, Jin S et al (2016) Multi-class multi-object tracking using changing point detection. In: European conference on computer vision (ECCV). Springer
24. Choi W (2015) Near-online multi-target tracking with aggregated local flow descriptor. In: 2015 IEEE international conference on computer vision (ICCV). IEEE Computer Society
25. Sanchez-Matilla R, Poiesi F, Cavallaro A (2016) Online multi-target tracking with strong and weak detections. In: Computer vision—ECCV 2016 workshops. Springer International Publishing
26. Bewley A, Ge Z, Ott L et al (2016) Simple online and realtime tracking. In: 2016 IEEE International Conference on Image Processing (ICIP). IEEE

Study on Routing Optimization Model of Container Sea-Rail Intermodal Transport Based on Transit Period

JunXiao Liu

Abstract Container sea-rail intermodal transport is one of the organizational forms of multimodal transport. It refers to providing an integrated container transport service by the effective convergence of water and railway transportation. It has long industry chain, high resource utilization, and good comprehensive benefits. On the basis of considering transit period, this paper proposes a route optimization model of container sea-rail intermodal transport by taking the minimization of total cost as the objective function. At the same time, this paper increases the constraints on transport capacity and transfer capacity. Finally, combining with a numerical example, the model is solved by optimization software. A container transport scheme with relatively least total cost meeting the transit period limit is acquired, which proves the feasibility of the model.

Keywords Container · Sea-rail intermodal transport · Route optimization · Transit period · LINGO

1 Introduction

Container sea-rail intermodal transport is an important mode of international multimodal transport in the word, which has outstanding advantages such as high efficiency, safety, and greenness. With the construction of China's railway passenger dedicated line and the continuous improvement of the railway network structure, the railway passenger and freight diversion will gradually take shape and the railway transport capacity will be greatly released, which creating favorable conditions for container sea-rail intermodal transport. In the sea-rail intermodal transport mode, goods are transported by trains to the ports directly, reloaded on ships and then transported to destinations, or unloaded from ships to ports and transferred by trains to destinations. It can achieve "one declaration, one inspection and one release" for the entire journey, which greatly improve the logistics efficiency. Sea-rail intermodal

J. Liu (✉)
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China
e-mail: 17120834@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_66

transport, especially the development of container sea-rail intermodal transport, combines the advantage of railway and water transport, and brings it into full play. It meets the needs of China's economic and social development and is an objective requirement for building an integrated transport system and a modern logistics system with complementary advantages and reasonable layout.

In order to further conform to the trend of the logistics industry, the scale of container sea-rail intermodal transport has been continuously expanding and the pace has been accelerating. According to data shows that, in the more developed western countries, the proportion of sea-rail intermodal transport in railway freight is generally between 20 and 30%, some regions can even be as high as 40–50%. In comparison, the proportion of China's sea-rail intermodal transport is very low, almost never exceeding 5%. Highway transport is still the main distribution method for port containers, and the proportion of highway transport is as high as 84%. This shows that the development of China's sea-rail intermodal transport is lagging behind, but from another point of view also shows that container sea-rail intermodal transport has greater development space. Moreover, although the scale of sea-rail intermodal transport is still small, the growth rate of container freight volume by railway transport has been significantly higher than the throughput growth rate.

2 Literature Review

At present, scholars at home and abroad have done a lot of research on container sea-rail intermodal transport and multimodal transport. Fang [1] pointed out that the low efficiency of railway freight transport organization and the lack of smooth sea-rail connection are the main reasons for the slow development of China's container sea-rail intermodal transport. The main reasons affecting the efficiency of railway container transport include: longer operating time of loading and unloading stations; longer transfer time of railway container trains; low running velocity of railway container trains; relatively complicated process of railway container operations and low punctuality of trains. And there are two main reasons for the lack of smooth sea-rail connection, one is the lack of connecting conditions between railway container yards and ports, the other one is incomplete construction of sea-rail intermodal transport information system.

Lozano [2] studied the shortest viable path problem in multimodal transport and presented an ad hoc modification of the Chronological Algorithm to solve the problem. He also considered the users' limitation on the number of modal transfers in order to obtain the corresponding viable solution set. Resat [3] considering time and congestion dependent vehicle speeds presented a multi-objective optimization model for integrating different transportation modes and used the augmented $\in$-constraint method minimize transportation cost and time simultaneously. Zhang [4] transformed the best route selection problem into a bi-objective shortest path problem through constructing the virtual network, developed a bi-objective integer programming model and solved the problem by a bi-objective labeling algorithm. Liang [5]

proposed a linear programming model to optimize vehicle routing in multimodal transport with the lowest cost and highest customer satisfaction as the optimization goal. She set customer satisfaction based on the customer's requirements for freight transportation time. Zhang [6] took the minimization of system cost as the objective function and proposed an optimal distribution model for multimodal transport networks. Liu [7] provided a dynamic ant colony algorithm to increase the ability of seeking optimal results and took the case of Tianjin Port to Mexico City as an example to prove the feasibility of the model. Li [8] was based on the comprehensive analysis of the current situation of container transportation and constructed a multi objective optimization model considering three aspects of transportation time, transportation costs and carbon emissions. Xie [9] took the user-defined time as an important constraint, established a multimodal transport path optimization model with a minimum total transportation cost objective and solved it with cross-entropy method.

At present, incomplete seamless connection between sea-rail intermodal transport, unstable punctuality of railway transport, and the high transfer cost have always restricted the development of China's sea-rail intermodal transport. In order to improve the level of railway and port container intermodal transport in China, this paper, based on the analysis of the development status of the domestic container sea-rail intermodal transport, establishes route optimization model of container sea-rail intermodal transport. We hope to contribute to the development of the logistics industry through this study.

3 Establish Model

In this section, we will establish a routing optimization model of container sea-rail intermodal transport aimed at minimizing the total cost during freight transportation.

3.1 Problem Description

In order to study the routing optimization model of container sea-rail intermodal transport, this paper uses diagram with multiple edges to describe the sea-rail intermodal network. Nodes are used to represent the origins, destinations and transfer nodes where the goods will be passed in transit. The connection between the nodes indicates the route of the cargo transportation. Different types of line segments between adjacent nodes represent different modes of transport. The adjacent nodes are two-way transportation, that is, the goods can be transported by trains from node a to c, and from node c to a there also can use railway transport to ship goods. The diagram is as shown in Fig. 1. For example, the goods can only be transported by railway or highway transport between node a and c.

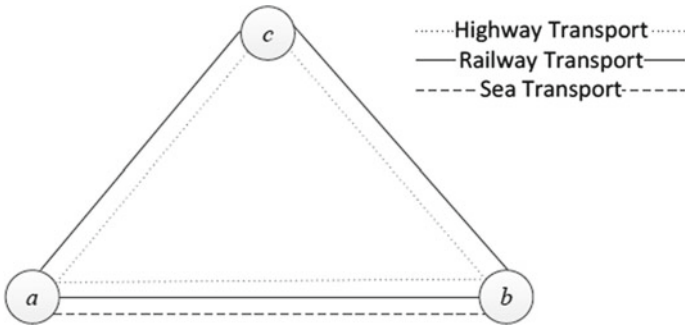


Fig. 1 The diagram with multiple edges

3.2 Assumptions

The following assumptions are made:

(1) Goods are inseparable during the transportation, that is, only a single mode of transport can be used between adjacent nodes;
(2) Different modes of transport between adjacent nodes have different distances;
(3) The transfer can only occur one time at most at each node;
(4) Seamless connection can be achieved between various modes of transport;
(5) This paper only considers transportation and transfer cost, and other costs are ignored;
(6) The parameters of each node are known.

3.3 Parameters and Variables

The following parameters are defined. These parameters are regarded as given conditions.

q Freight volume in the sea-rail intermodal transport network;
$a_{i,j}^{k}$ Unit transportation cost when using transport method k shipping goods from node i to node j;
c_{i}^{kl} Unit transfer cost when transport method changing from k to l at node i;
$Q_{i,j}^{k}$ Freight capacity when using transport method k shipping goods from node i to node j;
f_{i}^{kl} Transfer capacity when transport method changing from k to l at node i;
T Transit period, that is, the latest time to arrive the destination;
$d_{i,j}^{k}$ The distance when using transport method k shipping goods from node i to node j;
v^{k} Average travel velocity of transport method k;
h_{i}^{kl}: Transfer time when transport method changing from k to l at node i.

Two decision variables are defined:

$x_{i,j}^{k}$ The value is 1 if the goods are shipped from node i to node j by transport method k and i is not equal to j. Otherwise, it is zero.

y_{i}^{kl} The value is 1 if the transport method changing from k to l at node i and k is not equal to j. Otherwise, it is zero.

3.4 Mathematical Formulation

The objective is to minimize the total cost in transit, which includes transportation cost and transfer cost. We use Z to indicate the total cost. Then we have the objective function as below:

$$\min Z = \sum_{i}\sum_{j}\sum_{k} q \cdot x_{i,j}^{k} a_{i,j}^{k} + \sum_{i}\sum_{k}\sum_{l} q \cdot y_{i}^{kl} c_{i}^{kl} \tag{1}$$

The constraints of this model as follows:

(1) Only one transport mode can be used between adjacent nodes.

$$\sum_{k} x_{i,j}^{k} = 1 \quad \forall i, j \in n, i \neq j, \forall k \in m \tag{2}$$

(2) The transfer can only occur one time at most at each node.

$$\sum_{k}\sum_{l} y_{i}^{kl} \leq 1 \quad \forall i \in n, \forall k, l \in m, k \neq l \tag{3}$$

(3) If there has a change of transport method from k to l at node i, the transport mode k is used to ship goods from the previous node h to node i and the transport mode l is used to ship goods from node i to the next node j.

$$\sum_{h} x_{h,i}^{k} = \sum_{l} y_{i}^{kl} \quad \forall h, i \in n, h \neq i, \forall k, l \in m, k \neq l \tag{4}$$

$$\sum_{j} x_{i,j}^{l} = \sum_{k} y_{i}^{kl} \quad \forall i, j \in n, i \neq j, \forall k, l \in m, k \neq l \tag{5}$$

(4) The freight flow volume between adjacent nodes cannot exceed the transport capacity of the selected transport mode.

$$x_{i,j}^{k} \cdot q \leq Q_{i,j}^{k} \quad \forall i, j \in n, i \neq j, \forall k \in m \tag{6}$$

(5) If goods transfer at the node, the freight flow volume needs to meet node's transfer capacity constraint.

$$y_i^{kl} \cdot q \le f_i^{kl} \quad \forall i \in n, \forall k, l \in m, k \ne l \tag{7}$$

(6) The inflow and outflow of goods is equal at each transfer node.

$$\sum_k x_{h,i}^k = \sum_l x_{i,j}^l \quad \forall h, i, j \in n, h \ne i \ne j, \forall k, l \in m, k \ne l \tag{8}$$

(7) The total transit time is less than or equal to the transit period.

$$\sum_i \sum_j \sum_k x_{i,j}^k \frac{d_{i,j}^k}{v^k} + \sum_i \sum_k \sum_l y_i^{kl} t_i^{kl} \le T$$
$$\forall i, j \in n, i \ne j, \forall k, l \in m, k \ne l \tag{9}$$

4 Case Analysis

In order to verify the validity of the model, a container multimodal transportation network with 14 nodes is constructed. The transportation network is shown in Fig. 2, in which node 1 is the origin node and node 14 is the destination node. Different types of connecting lines between adjacent nodes represent different modes of transport.

The distance and freight capacity among different modes of transport are listed in Table 1 ("—" indicates that there is no such mode of transport between adjacent nodes, similarly hereinafter). Table 2 shows transfer capacity among different modes of transportation. In Table 3, transfer cost and time among different modes of transportation are listed.

Assume that there have 40 international standard containers of goods need to be shipped from node 1 to node 14, railway freight 3 yuan/(TEU km), highway freight is 8 yuan/(TEU km), and sea freight is 950 yuan/TEU. In this case, the average travel

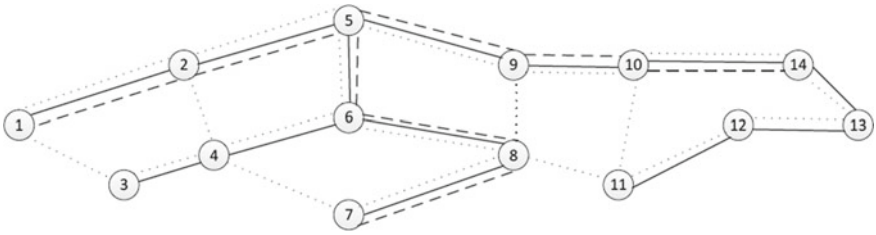


Fig. 2 A container multimodal transportation network

Table 1 The distance and freight capacity among different modes of transport

Node	Distance (km)			Freight capacity (TEU)		
	Railway transport	Highway transport	Sea transport	Railway transport	Highway transport	Sea transport
1–2	589	568	597	47	59	78
1–3	–	316	–	–	67	–
2–4	–	360	–	–	69	–
2–5	292	364	360	65	38	36
3–4	376	404	–	80	55	–
4–6	544	500	–	43	58	–
5–6	360	412	444	55	61	38
5–9	292	377	368	57	62	35
6–8	584	572	584	70	77	53
7–8	596	540	560	78	68	47
8–9	–	576	–	–	48	–
8–11	–	528	–	–	53	–
9–10	388	440	476	71	64	32
10–11	–	260	–	–	41	–
10–14	730	724	903	61	54	55
11–12	608	602	–	49	43	–
12–13	416	440	–	55	38	–
13–14	336	468	–	69	60	–

Table 2 Transfer capacity among different modes of transport

Node	Railway-highway transport	Railway-sea transport	Highway-sea transport
2	36	38	63
3	64	–	–
4	68	–	–
5	69	67	60
6	57	63	71
7	75	68	40
8	76	53	48
9	74	39	67
10	68	43	65
11	64	–	–
12	67	68	76
13	72	58	38

Unit TEU

Table 3 Transfer cost and time among different modes of transport

The mode of transport	Transfer cost (yuan/TEU)	Transfer time (h/TEU)
Railway-highway transport	5	0.07
Railway-sea transport	7	0.15
Highway-sea transport	10	0.1

velocity of railway transport is 50 km/h, the average travel velocity of highway transport is 60 km/h, and the travel speed of sea transport is 30 km/h. The transit period of goods is limited to 3 days.

After solving the model with commercial optimization software LINGO 11.0, the optimal solution is 5913 yuan, and the total transportation time is 67.32 h. The transport scheme is as follow: starting from node 1, passing through node 2, 5, 9 and 10 by railway transport, transferring sea transport at node 10, shipping from node 10 to node 14 by sea transport.

5 Conclusions

Container sea-rail intermodal transport has become the priority mode of transport for all countries in the world by its unique advantages, such as high transport capacity, low transportation cost, high transport safety, and low pollution emission. This paper establishes a route optimization model of container sea-rail intermodal transport with the goal of minimizing total cost by taking transit period as a constraint. It not only meets the customer's requirements for transportation efficiency, but also reduces the logistics transportation cost. A more reasonable container transport scheme can be obtained through this paper. Separately pursuing the minimization of transport time or transport cost may lead to the loss of source of goods and reduce the proportion of container sea-rail intermodal transport in the transportation market. This paper optimizes the route of sea-rail intermodal transport on the basis of the transit period, which can greatly improve the customer satisfaction. Therefore, this paper has a certain practical significance.

In this article, the values of some parameters are selected based on the average value of those in the daily transportation process. For example, transportation cost and transfer cost may change with the freight volume, however, they are constants in this paper. That may be some deviation from the actual transport situation, so there still needs to be further studied.

References

1. Fang QG (2016) Development strategies of rail-water container intermodal transportation. J Transp Syst Eng Inf Technol 16:31–36
2. Lozano A, Storchi G (2001) Shortest viable path algorithm in multimodal networks. Transp Res Part A 35:225–241
3. Resat HG, Turkay M (2015) Design and operation of intermodal transportation network in the Marmara region of Turkey. Transp Res Part E Logistics Transp Rev 83:16–33
4. Zhang Y, Wang YX, Yang HL (2013) Best routes selection in international intermodal networks. J Dalian Marit Univ 39:111–114
5. Liang XK (2017) A study on optimizing routes of railway containerization multimodal transport with fuzzy time constraints. Railway Transp Econ 39:55–60
6. Zhang JY, Guo YH (2002) Multimode transportation network assignment model. J China Railway Soc 24:114–116
7. Liu WL (2012) Route optimization of international multimodal container transportation based on dynamic ant colony algorithm. J Beijing Jiaotong University (Social Sciences Edition) 11:57–62
8. Li YM, Guo XY, Yang L (2017) Route optimization of China-EU container multimodal transport considering various factors. J Railway Sci Eng 14:2239–2248
9. Xie T, Yi M, Liu CN (2015) Study on routing optimization model of inter-modal transportation with time limit. Logistics Eng Manage 37:56–58

Passenger Flow Prediction Model of Intercity Railway Based on G-BP Network

Hai-lian Li, Meng-kai Lin and Qi-cai Wang

Abstract Inter-city railway as the city's comprehensive transportation system, the development of urban industrial economy and the image of the overall improve greatly boost. However, scientific and reasonable forecast traffic is the focus on the study of the inter-city railway construction project, which aim is to obtain the characteristics and rules of passenger flow, planning area to provide comprehensive system for railway planning and the actual resources and foundation of real and reliable data. Based on the grey relational analysis method influence the traffic data and the relationship between influencing factors, choose the main influence factors of traffic influence factors of the BP neural network model is established. Finally combined Lanzhou to Zhongchuan Airport inter-city railway project to traffic prediction research and survey data, it is concluded that the influence factors of the BP neural network model has good predictability to the traffic.

Keywords Passenger traffic volume · Prediction · Grey theory · BP neural network

1 Introduction

As an important public welfare infrastructure for urban construction, the intercity railway is a key project with large investment, long construction period and comprehensive comprehensiveness. Once it was completed, it will be difficult to change.

H. Li (✉) · M. Lin · Q. Wang
Key Laboratory of Road & Bridge and Underground Engineering of Gansu Province, Lanzhou Jiaotong University, Lanzhou, China
e-mail: 17834040@qq.com

M. Lin
e-mail: 27834040@qq.com

Q. Wang
e-mail: 13909486262@139.com

National and Provincial Joint Engineering Laboratory of Road & Bridge Disaster Prevention and Control, Lanzhou Jiaotong University, Lanzhou, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_67

Therefore, it is necessary to prepare adequately in early stage of project construction, and one of the key points of feasibility study phase in early stage project is to predict passenger flow of the pre-built intercity railway [1], and only guarantee accurate forecasting of passenger flow data can enable the newly built intercity railway to achieve the expected economic and social benefits. Nowadays, the forecast of passenger traffic is no longer a simple linear change of time series. It is affected by the socio-economic level, the population and living standards, the national economic system and policies. Meanwhile, various modes of transportation also influence and restrict passenger flow data on the same interval line [2]. It is impossible to properly deploy the transportation resources of the transportation market without a scientific and reasonable method for passenger flow forecasting [3].

Based on economic development level of intercity railways, development of population structure, living standards of urban residents, and development of tourism industry, this paper constructs a passenger flow forecasting model based on G-BP network with establishing an inter-city railway passenger flow forecasting index system and combining with grey correlation analysis and BP. And further analyzing influence mechanism of each factor to realize the quantitative prediction and qualitative analysis of intercity railway passenger flow.

2 Factors Affecting Passenger Flow on Intercity Railway

The analysis of intercity railway passenger flow is a complex system engineering, involving multiple visible factors and invisible factors. With development of urban regional economy and progress of urbanization, how to improve transportation capacity, ease traffic pressure, and ensure unimpeded travel is a bottleneck restricting the development of each city. The construction of intercity railways and improvement of intercity passenger transport are main measures to solve this bottleneck. Intercity railway passenger flow demand is main factor affecting the analysis of intercity railway passenger flow [4]. Meanwhile, level of regional economic development, degree of urbanization, scale and layout of urban development, and level of other transportation development within the adjacent cities have a direct impact on the analysis of intercity railway passenger flow. In addition, intercity railway passenger flow is also affected by hidden factors such as population structure and development of neighboring cities, urban residents' living standards, urban resource environment, passenger service level, and tourism industry development. Establishing intercity railway passenger flow forecasting index system are shown in Table 1 by comprehensively analyzing interaction and dependency relationship between these influencing factors.

Table 1 Intercity railway passenger flow forecasting index system

Destination layer	Intercity railway passenger flow			
Criterion layer	Dominant factor			Recessive factor
	Passenger demand	Regional economy	Regional traffic	Environment and service
Indicator layer	Traffic division	Gross product	Traffic mode	Population
	Total traffic	Gross value of the primary industry	Travel convenience	Population structure
	Regional per capita travel times	Total value of the secondary industry	Traffic throughput	Residents' living standards
	Floating population	Gross value of the tertiary industry	Regional traffic connection	Urban resource environment
	Travel release measures	Per capita disposable		Passenger service
	Urbanization	Consumer price index level		Tourism development

3 Intercity Railway Passenger Flow Forecasting Model

According to analysis index system of intercity railway passenger flow, the grey relational analysis and BP neural network theory are used to establish the intercity railway passenger flow forecasting model which achieves the scientific evaluation and analysis of urban regional intercity railway passenger flow.

3.1 Prediction Theory

Grey relational analysis as a technical method of grey theory is widely used [5, 6]. It has no special requirements on the amount of sample data and the regularity of sample, and it's not computationally intensive and very convenient. The grey correlation analysis is based on sample data and gray correlation degree to determine importance degree of each factor [7]. Therefore, this paper uses the grey correlation analysis method to calculate the correlation coefficient between each index, so as to optimize the intercity railway passenger flow forecasting index system and extract the main forecasting indicators. At the same time, according to main indicators of intercity railway passenger flow forecast, corresponding BP model is established, that is, main indicators are used as the network input, and traffic data is used as the output to find out internal relationship between traffic data and variation index, then predict the future passenger flow.

3.2 Prediction Steps

Based on above theoretical analysis, the specific prediction steps are as follows:

1. Extraction of passenger traffic forecast main indicators for intercity railways [8, 9].

(1) Nondimensionalization of data.

Setting the original data sequence of passenger flow prediction index of intercity railway as $y_i = (y_i(1), y_i(2), y_i(3), \ldots, y_i(n))$, dimensionless data sequence after processing is:

$$x_i = \left(\frac{y_i(1)}{y_i(1)}, \frac{y_i(2)}{y_i(1)}, \frac{y_i(3)}{y_i(1)}, \ldots, \frac{y_i(\mathrm{n})}{y_i(1)}\right)$$
$$(i = 1, 2, 3 \ldots, \mathrm{n}, \quad y_i(1) \neq 0) \tag{1}$$

(2) Evaluation of sequence difference. That is absolute difference between x_i and x_0 corresponding to each point.

$$\Delta_i = |x_0(k) - x_i(k)|, (i, k = 1, 2, \ldots, n) \tag{2}$$

(3) Calculating two-level maximum difference and minimum difference.

Two-level maximum difference:

$$M = \max_i \max_k \Delta_i(k) \tag{3}$$

Two-level minimum difference:

$$m = \min_i \min_k \Delta_i(k) \tag{4}$$

(4) Calculating correlation coefficient.

$$\gamma_{0i}(k) = \frac{\min\limits_i \min\limits_k |x_0(k) - x_i(k)| + \zeta \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \zeta \max\limits_i \max\limits_k |x_0(k) - x_i(k)|},$$
$$\zeta \in (0, 1), k = 1, 2, \ldots, n; \quad i = 1, 2, \ldots, m \tag{5}$$

In this formula: $\gamma_{0i}(k)$ is the relative difference between comparison curve x_1 and the reference curve x_0 at k-th time, which means correlation coefficient of x_1 versus x_0 at time k. ζ is resolution coefficient, we take 0.5 in this article.

(5) Calculating degree of relevance, and take average of each correlation coefficient in each group as association degree.

$$\gamma_{0i} = \frac{1}{n}\sum_{k=1}^{n}\gamma_{0i}(k), \quad i = 1, 2, \ldots, m \tag{6}$$

(6) According to the order of prediction index correlation ranking, choose larger relevance index as main indicator of prediction.

2. Establish a predictive index BP network. Main indicator is used as network input, and traffic data is used as output to complete the forward calculation and error back propagation process of network [10–12]. The specific process is as follows.

(1) Forward propagation of signals.

Input of i-th node of the hidden layer net_i is:

$$net_i = \sum_{j=1}^{M} w_{ij}x_j + \theta_i \tag{7}$$

Output of i-th node of the hidden layer o_i is:

$$o_i = \varphi(net_i) = \varphi\left(\sum_{j=1}^{M} w_{ij}x_j + \theta_i\right) \tag{8}$$

(2) Backpropagation of errors. The quadratic error criterion function for each sample p is E_P.

$$E_P = \frac{1}{2}\sum_{k=1}^{L}(T_k - o_k)^2 \tag{9}$$

(3) Updating weight and threshold

$$w_{ij} = w_{ij} + \eta H_j\left(1 - H_j\right)x(i)\sum_{k=1}^{m} w_{jk}e_k,$$
$$i = 1, 2, \ldots, n; \ j = 1, 2, \ldots, l \tag{10}$$

$w_{jk} = w_{jk} + hH_j e_k$, $j = 1, 2, \ldots, l$; $k = 1, 2, \ldots, l$, $a_j = a_j + \eta H_j\left(1 - H_j\right)\sum_{k=1}^{m} w_{jk}e_k$, $j = 1, 2, \ldots, l$, $b_k = b_k + e_k$, $k = 1, 2, \ldots, m$, η is learning rate, e_k is prediction error.

(4) Determine whether the indicator satisfies accuracy requirement. If not, return to step 1 to repeat calculation.

3. According to the calculation results, obtain the predicted value of intercity railway passenger flow.

4 Instance Prediction

Using above predicted models and methods to predict passenger flow of the newly built Lanzhou-Zhongchuan intercity railway. The specific prediction process is as follows.

4.1 Predictive Indicators

Due to limitations of objective conditions and collection methods, some indicators must be reduced because they are default and ambiguous in the intercity railway passenger flow forecasting index system. This paper chooses 10 indicators such as total production value (100 million yuan), total value of primary industry (100 million yuan), total secondary industry(100 million yuan), total value of tertiary industry (100 million yuan), total population (10,000 people), total urban population (10,000 people), per capita disposable income (10,000 people), consumer price index level, transportation and communication expenses, and Zhongchuan Airport throughput (10,000 people) as predictive indicators from four aspects of Lanzhou's regional economic development level, population structure development, urban residents' living standards and tourism industry development. And combined with the above indicators data of Lanzhou from 2006 to 2014 as original data, the gray correlation analysis model was established to analyze passenger flow.

4.2 Main Indicators Extraction

Let x_0 represents passenger flow data, $x_1 \sim x_{10}$ represents total production value (100 million yuan), total value of primary industry (100 million yuan), total value of secondary industry (100 million yuan), total value of tertiary industry (100 million yuan), total population (10,000 people), towns total population (10,000 people), per capita disposable income (10,000 people), consumer price index level, transportation and communication expenses, and Zhongchuan Airport throughput (10,000 people) respectively. Using grey system modeling software to calculate the Dunsian correlation between the sequence $x_1 \sim x_{10}$ and traffic data.

1. Calculating initial value of sequence which means nondimensionalization of data.
   Data1 =

1.000	1.464	1.951	2.660	3.268	3.945	4.734	5.597	6.593
1.000	1.148	1.325	1.493	1.723	2.130	2.450	2.782	2.997
1.000	1.148	1.236	1.344	1.487	1.760	1.986	2.161	2.358
1.000	1.157	1.371	1.493	1.822	2.261	2.565	2.825	2.856
1.000	1.139	1.291	1.419	1.652	2.039	2.381	2.787	3.168
1.000	1.023	1.050	1.100	1.197	1.220	1.239	1.276	1.315
1.000	1.009	1.027	1.023	1.100	1.100	1.131	1.192	1.234
1.000	1.091	1.240	1.355	1.493	1.694	1.958	2.205	2.445
1.000	1.033	1.043	0.978	1.018	1.038	1.011	1.042	1.046
1.000	0.992	0.977	0.992	0.992	1.000	1.000	1.004	1.013
1.000	1.349	1.189	1.538	1.936	2.047	2.463	3.036	3.540

2. Calculating the difference sequence, that is absolute difference between x_1 and x_0 at each moment.
   Data2 =

0.000	0.317	0.625	1.167	1.544	1.815	2.283	2.815	3.596
0.000	0.316	0.715	1.316	1.781	2.185	2.748	3.436	4.234
0.000	0.307	0.579	1.167	1.445	1.684	2.169	2.772	3.737
0.000	0.325	0.660	1.241	1.616	1.906	2.353	2.810	3.425
0.000	0.441	0.901	1.560	2.070	2.725	3.494	4.321	5.278
0.000	0.455	0.924	1.637	2.168	2.845	3.603	4.405	5.359
0.000	0.374	0.711	1.305	1.774	2.251	2.775	3.392	4.147
0.000	0.432	0.907	1.682	2.250	2.907	3.723	4.555	5.546
0.000	0.472	0.974	1.668	2.276	2.945	3.734	4.593	5.580
0.000	0.115	0.762	1.122	1.331	1.898	2.271	2.561	3.052

3. Calculating the range. Looking for maximum and minimum values in Data2, you can get maximum difference of 5.580, and minimum difference is 0.000.

4. Calculating the correlation coefficient.
   Data3=

1.000	0.898	0.817	0.705	0.644	0.606	0.550	0.498	0.437
1.000	0.898	0.796	0.679	0.610	0.561	0.504	0.448	0.397
1.000	0.899	0.804	0.705	0.632	0.601	0.549	0.497	0.427
1.000	0.896	0.809	0.692	0.633	0.594	0.542	0.498	0.449
1.000	0.863	0.756	0.641	0.574	0.506	0.444	0.392	0.346
1.000	0.860	0.751	0.630	0.563	0.495	0.436	0.388	0.342
1.000	0.882	0.797	0.681	0.611	0.553	0.501	0.451	0.402
1.000	0.866	0.755	0.624	0.554	0.490	0.428	0.380	0.335
1.000	0.855	0.741	0.626	0.551	0.486	0.428	0.378	0.333
1.000	0.960	0.785	0.713	0.677	0.595	0.551	0.521	0.478

5. According to correlation degree formula, obtain gray correlation degrees are: 0.684, 0.655, 0.679, 0.679, 0.614, 0.607, 0.653, 0.603, 0.600, 0.698 between passenger flow and various influencing factors.
6. Main indicator determination. Through analysis the correlation data of ten forecast indicators in the four aspects of Lanzhou's regional economic development level, population structure development, urban residents' living standards and tourism industry development, select total production value (100 million yuan) and population of Lanzhou City (10,000 people), per capita income (yuan), and throughput of Zhongchuan Airport (10,000 people) as main indicators of passenger flow prediction. The population is proportional to passenger flow. The increase in per capita income increases willingness to travel. The gross domestic product reflects level of local economic development. And the airport throughput is positively correlated with passenger flow along the railway.

4.3 Passenger Flow Prediction

Let x_1, x_2, x_3, x_4 represents population, per capita income, gross domestic product, and airport throughput respectively, and as the input of BP neural network. Let y represents the annual passenger traffic data value and use it as BP neural network output. The BP neural network model was established by programming with Matlab7.0. The original values of population, per capita income, gross production value and airport throughput index of Lanzhou City from 2006 to 2014 are shown in Table 2.

1. Raw data preprocessing. In order to eliminate dimensional influence of various indicator data affecting passenger flow, input data and output data of BP

Table 2 Indicator data affecting passenger flow from 2006 to 2014

Years	Gross product (100 million yuan)	Population (10,000 people)	Residents' disposable income (yuan)	Zhongchuan Airport throughput (10,000 people)
2006	638.47	243.68	9418.00	186.11
2007	732.76	245.83	10,271.18	251.09
2008	846.28	250.15	11,676.77	221.23
2009	952.98	249.22	12,761.00	286.15
2010	1100.39	267.93	14,061.84	360.35
2011	1360.03	268.08	15,952.57	380.90
2012	1564.41	275.49	18,442.76	458.35
2013	1776.28	290.52	20,767.00	564.96
2014	1913.5	300.65	23,030.00	658.88

neural network are standardized to solve the comparability between data indicators. In prediction of this model, sample data is dimensionlessly processed using normalization function mapminmax in Matlab7.0.

2. Training BP network. The training function of BP neural network is newff, sim is a simulation function, the number of input layer nodes is n = 4, and the number of output layer nodes is m = 1, and the number of hidden layer nodes is calculated by empirical formula $n_1 = \sqrt{n+m} + a$ (a is a constant between [1, 10]), the final value is set to 6 after repeated precision trials. And use Matlab7.0 software to program simulating calculation.
3. Analysis of prediction results. Results are denormalized to obtain annual passenger flow prediction value, and compared with the Lanzhou-Zhongchuan passenger flow survey data from 2006 to 2014. The analysis results are shown in Table 3.

It can be seen from the test results in Table 3 that error between the model test results and real results is small, the average error is 3.37%, and the test is efficiency. That indicates the prediction model has achieved satisfactory results and feasible to passenger flow data. The construction and opening of the Lanzhou-Zhongchuan intercity railway not only ease a large number of passenger flow in airport and new district, but also reduce traffic pressure of the city, and greatly save travel time of passengers and promote area economic development along the route. The accurate forecasting results also provide an important basis for economic benefit analysis of intercity railway operation in later period.

Table 3 Prediction results of BP model from Lanzhou to Zhongchuan from 2006 to 2014 [10,000 people]

Years	2006	2007	2008	2009	2010	2011	2012	2013	2014
Actual value	168.26	246.37	328.24	447.56	549.80	663.80	796.48	941.76	1109.26
Predictive value	170.32	280.84	345.79	455.19	560.58	651.67	812.80	952.27	1121.62
Relative error	1.22	13.99	5.35	1.70	1.96	1.83	2.05	1.12	1.11
Average error	3.37								

5 Conclusions

Here we may draw the following conclusions.

(1) This paper further clarifies the meaning and scope of forecasting index system used in intercity railway passenger flow through in-depth analysis and research on influencing factors of intercity railway passenger flow forecasting and its linkages.
(2) Based on grey correlation method and BP neural network theory, the G-BP network-based intercity railway passenger flow prediction model and method are established. This model not only makes prediction process more scientific, intuitive and convenient, but also improves accuracy of prediction. Meanwhile, it provides more effective basic data and theoretical support for design and optimization of intercity railway lines in China in the future.

Acknowledgements The authors would like to acknowledge the financial support provided by the National Natural Science Foundation of China (Grant No. 51868042); the Changjiang Scholars and Innovation Team Development Program of Ministry of Education (IRT_15R29); Youth Science Foundation of Gansu (17JR5RA087); Youth Science Foundation of Lanzhou Jiaotong University (2017016); Foundation of A Hundred Young Talents Training Program of Lanzhou Jiaotong University (2018103).

References

1. Wang C (2008) Prediction of civil aviation passenger traffic volume on grey theory and RBF neural network. Beijing Jiaotong University
2. Liu Q (2008) Research on travel demand forecasting based on complex system. China Academy of Railway Sciences
3. Mou Z (2008) The prophecy and analyse about Shen Yang area's highway-flow. Northeastern University
4. Deng L (2009) Research on the influencing factors on intercity rail construction. J Railway Eng Soc 10:99–101
5. Liu S, Dang Y, Fang Z (2010) Grey system theory and its application, 3rd edn. Science Press, Beijing
6. Xian M (2016) The research of railway passenger flow forecasting method based on greyneural network models. Southwest Jiaotong University
7. Feng B, Bao X, Wang Q, Dong X (2015) A new combination model for forecasting railway passenger volume. Railway Stan Des 59(12):6–9
8. Feng B, Bao X, Wang Q (2015) Research of railway passenger volume forecast based on grey and neural network. J Railway Sci Eng 12(05):1227–1231
9. Ting G, Li W (2013) Prediction research of highway traveling passenger volume based on wavelet neural network. Appl Mech Mater 401–403:1401–1405
10. Meng T, Xiang F, Jing Z, Bin R (2014) FA-BP neural network-based forecast for railway passenger volume. Appl Mech Mater 641(8):673–677

11. Wang H, Wang X, Dang J (2010) Comprehensive evaluation of comfort of high-speed trains based on fuzzy reduction. J China Railway Soc 32(05):98–102
12. Wu H, Zhen J, Wang Y, Wang F (2014) Railway passenger and freight prediction based on RBF neural network theory. J Railway Sci Eng 11(4):110–113

Comparative Study on Value of a Statistical Life in Road Traffic Based on Mixed Logit Model

Wen-ge Liu and Sheng-chuan Zhao

Abstract In order to further improve the accuracy of evaluation models on value of a statistical life (VOSL) in road traffic, four mixed logit (ML) models of route-choice with truncated normal distribution and lognormal distribution were used to construct VOSL models. A route-choice questionnaire was designed by the stated choice method, and the traffic survey was carried out in Dalian with survey data obtained. Monte Carlo simulation algorithm was used to calibrate parameters by 150 simulations, and the 4 ML models were analyzed comparatively. Finally, the VOSL estimate of private drivers in Dalian and its distribution function were obtained. The research results indicate: ML model with truncated normal distribution has $\overline{\rho}^2$ of 0.1516 and hit ratio of 70.42%, which has a lower accuracy. 3 ML models with lognormal distribution have a high accuracy, whose $\overline{\rho}^2$ are all between (0.2–0.4) and hit ratios all above 80%. The 4th ML model whose parameters of fatal risk and travel cost obeying lognormal distribution simultaneously has the highest accuracy, with the greatest $\overline{\rho}^2$ (0.2534) and the highest hit ratio (84.76%). VOSL in road traffic based on the 4th ML model obeys lognormal distribution with parameters $(2.0622, 0.6740^2)$ with the mathematical expectation of 986,840 RMB. The maximum probability is 9.45% when VOSL is 500,000 RMB.

Keywords Traffic safety · Value of a statistical life · Stated choice method · Mixed logit model · Monte Carlo simulation

1 Introduction

The conception of value of a statistical life (VOSL) is first proposed by Schelling (1968), which means the average of individual marginal rate of substitution of the willingness to pay (WTP) for reduced fatal risks or the willingness to accept for

W. Liu (✉)
School of Economics and Management, Dalian Jiaotong University, Dalian 116028, China
e-mail: 18737653@qq.com

S. Zhao
School of Transportation and Logistics, Dalian University of Technology, Dalian 116024, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_68

increased fatal risks from a sample [1]. Using VOSL to measure the value of human life has been recognized by most economists. VOSL has been widely studied in many safety fields such as environmental pollution, food hygiene, occupational safety and traffic safety. VOSL in road traffic has become an indispensably basic indicator for the cost-benefit analysis on the traffic safety project.

According to statistics from the Ministry of Public Security, there were 213,000 road traffic accidents in China in 2016, resulting in 63,093 people died. In terms of the death rate of 10,000 vehicles in 2011, the figures are between 0.7 and 2 in Japan, Europe and America, but 6.2 in China. According to relevant data of Road Traffic Safety Development Report (2017), the death rate of 10,000 vehicles in China is 2.14, which is still higher than that in other countries. Therefore, it is of great significance to study VOSL in road traffic in China.

The evaluation methods of VOSL in road traffic are mainly based on the revealed preference method (RP) and the stated preference method (SP). RP includes wage-risk method and hedonic pricing method, which have been paid little attentions because of great limitations [2]. SP includes the contingent value method (CVM) and the stated choice method (SCM). CVM is most widely used, but it has two main limitations that are the hypothesis bias and the difficulty in risk communication [3]. SCM is a relatively new method, but it has good research prospects [4] because it can effectively compensate for the limitations of CVM.

At present, the evaluation models of VOSL based on SCM mainly include binary logit (BL) model and mixed logit (ML) model. Rizzi and Ortúzar (2003) established VOSL model based on BL model, and obtained VOSL estimates of round-trip population between San Diego and Valparaiso between $149,000 and $206,000 [5]. Iragüen and Ortúzar (2004) established VOSL model based on BL model, and obtained VOSL estimates of Santiago residents as $125,000 [6]. Zhao and Zhang (2008) proposed the theoretical derivation of VOSL based on BL model [7]. The author (2013) established a route-choice model based on BL model, and obtained VOSL estimates of private drivers in Dalian as RMB 360,000 [8] BL model is simple to construct and easy to solve, but it has the characteristic defect of IIA [9]. Moreover, VOSL estimates are fixed values based on BL model, which cannot objectively reflect the changing regulation of VOSL.

ML model can overcome the IIA defect and its parameters are not fixed values but random distribution, so it has high flexibility [10]. Besides, on the premise of guaranteeing samples, ML model has sufficient reliability and strong ability to reproduce original data. In recent ten years, ML model has been gradually concerned and applied in VOSL evaluation. Rouwendal et al. (2009) constructed VOSL model based on ML model with uniform distribution, triangular distribution and lognormal distribution [11]. Rheinberger (2011) evaluated VOSL of Alpine road travelers based on ML model with triangular distribution, and the VOSL estimates were in the range of CHF 6.0–7.8 million [12]. The author (2013) evaluated VOSL of private drivers in Dalian based on ML model with lognormal distribution, and the VOSL estimates were in the range of RMB 350,000–370,000 [13]. Yang et al. (2016) evaluated VOSL of motorists and non-motorists in Nanjing based on ML model with normal distribution and lognormal distribution, and the VOSL estimates were separately RMB

3.72 million and RMB 3.28 million [14]. In summary, the research results indicate that: (1) ML models with uniform and triangular parameters are less accurate; (2) ML model with normal parameters is more accurate than the former. However, the coexistence of positive and negative values of its parameters is inconsistent with the one-way preference of model variables. Therefore, its evaluation results have larger errors; (3) ML model with lognormal parameters is more accurate than the former. Moreover, its parameters are positive, which accords with the one-way preference of model variables. However, there are insufficient explanations for other key variables except for the fatal risk variable.

In view of the above, in order to further improve the accuracy of VOSL estimates and explain its distribution regulation more reasonably, a ML model based on truncated normal distribution and three ML models based on lognormal distribution are proposed and compared. There are three research purposes: (1) Constructing VOSL evaluation models based on four ML models with truncated normal and lognormal distribution; (2) Designing SCM questionnaire and conducting traffic survey in Dalian to obtain survey data. (3) Calibrating parameters of VOSL evaluation models by Monte Carlo simulation algorithm and making a comparative study among the four ML models, with VOSL estimates and its distribution rule obtained finally.

2 SCM Questionnaire Design

SCM provides a hypothetical scenario with two or more alternatives and respondents are asked to choose a preferred one from them. SCM is based on random utility theory, whose basic assumption is that policy-makers pursue utility maximization. In order to improve investigation efficiency, the orthogonal experiment method is combined with SCM to design the questionnaire. Questionnaire design is divided into three steps. Firstly, the hypothetical scenario of route-choice is designed. Then the characteristic variables are chosen and variables levels are determined. Finally, the combination of alternatives is designed and the questionnaire items are determined.

2.1 Hypothetical Scenario

"Dalian-Lvshun route choice" is taken as the hypothesis scenario. That is to say, there is a toll road and a free road between Dalian and Lvshun. The fatal risk of the toll road is relatively low, and that of the free road is relatively high. The travel times of two roads are different. Private car commuters between Dalian and Lvshun are taken as respondents, and asked to choose a preferred road between two roads.

Table 1 Levels of characteristic variables

Variables Levels number	A toll road			A fee road	
	Fee (RMB)	Travel time (minutes)	Fatal risk (yearly deaths/10^5 persons)	Travel time (minutes)	Fatal risk (yearly deaths/10^5 persons)
1	20	40	2/100,000	60	8/100,000
2	15	50	5/100,000	70	12/100,000
3	10	60	8/100,000	80	15/100,000

2.2 Characteristic Variables and Variables Levels

Three variables of travel cost, travel time and fatal risk are selected as characteristic variables. According to relevant data from Dalian Public Security Bureau, the average fatal risk of drivers in Dalian in 2016 is 8/100,000. The distance between Dalian and Lvshun is about 50 km with average travel time of 60 min. According to the charge standard of expressways and the economic development in Dalian, the average charge is 15 RMB per time. According to the average value of variables, variables are set to three levels (shown in Table 1).

2.3 Alternative Combinations

As seen in Table 1, the number of alternative combinations is $3^5 = 243$ if the comprehensive experimental method is adopted. The orthogonal experiment method can greatly reduce experimental numbers and improve investigation efficiency. In this study, the orthogonal experiment method was used and the number of alternative combinations was reduced to $3^{5-2} = 27$. Finally, 27 alternative combinations were randomly sorted and divided into 9 questionnaires, each of which contained three alternative combinations.

3 Random Distributions and VOSL Evaluation Model

3.1 Random Distributions and Their Characteristics

(1) Truncated normal distribution

When random variable x obeys normal distribution $x \sim N(\mu, \sigma^2)$, its density function is expressed as:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-u)^2}{2\sigma^2}} \tag{1}$$

Let $y = \max(0, x)$, then random variable y obeys truncated normal distribution. Truncated normal distribution curve is only on the right side of zero for normal distribution, which overcomes the defect that normal distribution curve has distribution around zero value. Therefore, it is suitable for parameter estimation of one-way preference factors [15].

(2) Lognormal distribution

When random variable x obeys normal distribution $x \sim N(\mu, \sigma^2)$, Let $y = \exp(x)$, then random variable y obeys lognormal distribution and its density function is expressed as:

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{(\ln y-u)^2}{2\sigma^2}}, (y > 0) \tag{2}$$

Lognormal distribution curve is only on the right side of zero, which is also suitable for parameter estimation of one-way preference factors.

Because the three characteristic variables of fatal risk, travel time and travel cost all have one-way preference characteristic, the above two random distributions are both suitable as the random distribution functions of their parameters.

3.2 VOSL Evaluation Model

In the route-choice scenario, assuming that the utility function for the respondent n to choose a route has a linear relationship with travel cost c, travel time t and fatal risk r, the fixed item V_{1n} of utility function for the respondent n to choose a toll road is expressed as:

$$V_{1n} = a_0 + \alpha c_{1n} + \beta t_{1n} + \chi r_{1n} \tag{3}$$

The fixed item V_{2n} of utility function for the respondent n to choose a free road is expressed as:

$$V_{2n} = \beta t_{2n} + \chi r_{2n} \tag{4}$$

According to the conception of VOSL, VOSL in road traffic can be deduced as follows:

$$VOSL = \frac{\partial V_n / \partial r}{\partial V_n / \partial c} = \frac{\chi}{\alpha} \tag{5}$$

In the formula above, N represents sample size.

Table 2 Variables and parameters of ML models

Variables	Parameters	Variable symbol	Variable description
Constant	a_0		
Travel cost	α	c	RMB
Travel time	β	t	Minutes
Fatal risk	μ	r	(deaths/10^5 persons)
Fatal risk variance	σ	r_s	(deaths/10^5 persons)

Because VOSL mainly analyses the differential impact of fatal risk on different travelers, it is firstly assumed that the fatal risk coefficient obeys the random distribution, and the coefficients of travel time and travel cost are constants.

Specific parameter assumptions are as follows: ① The higher the fatal risk faced by travelers, the less the utility is for choosing the route. Therefore, the coefficient of fatal risk is negative. Meanwhile, it is assumed that the probability is small when fatal risk coefficient is very large or very small, and it is the highest when in the middle level. Therefore, the fatal risk coefficient χ is assumed to obey truncated normal and lognormal distribution. ② The higher the travel cost faced by travelers, the less the utility is for choosing the route. Meanwhile, assuming travelers' endurance of payment is the same, so the travel cost coefficient α is assumed to be a negative constant. ③ The longer the travel time faced by travelers, the less the utility is for choosing the route. Meanwhile, assuming travelers' endurance of travel time is the same, so its coefficient β is assumed to be a negative constant.

According that the fatal risk coefficient obeys truncated normal and lognormal distribution, two kinds of ML models are constructed. Model variables and parameters are shown in Table 2, and utility functions of ML models are shown in Table 3.

How to obtain coefficients of fatal risk and travel cost is the key to determine VOSL. The parameters of ML models can be calibrated by Monte Carlo simulation algorithm.

4 Monte Carlo Simulation Algorithms

4.1 Selection Probability of ML Model

The probability P_{1n} of the respondent n choosing a toll road can be expressed as:

$$P_{1n} = \int \frac{e^{V_{1n}(\alpha,\beta,\chi)}}{\sum_{j=1}^{2} e^{V_{jn}(\alpha,\beta,\chi)}} f(\chi) d\chi \tag{6}$$

Table 3 Utility functions of ML models with the fatal risk parameter obeying different distributions

ML models	Hypothetical condition	Fixed items of utility function of ML models
ML-1	① Fatal risk coefficient χ obeys negative truncated normal distribution ② Travel time coefficient β is a negative constant ③ Travel cost coefficient α is a negative constant	$V_{1n} = a_0 + \alpha c_{1n} + \beta t_{1n} - \max(0, \mu + \varepsilon\sigma) r_{1n}$ $V_{2n} = \beta t_{2n} - \max(0, \mu + \varepsilon\sigma) r_{2n}$
ML-2	① Fatal risk coefficient χ obeys negative lognormal distribution ② Travel time coefficient β is a negative constant ③ Travel cost coefficient α is a negative constant	$V_{1n} = a_0 + \alpha c_{1n} + \beta t_{1n} - \exp(\mu + \varepsilon\sigma) r_{1n}$ $V_{2n} = \beta t_{2n} - \exp(\mu + \varepsilon\sigma) r_{2n}$

Note μ and σ are respectively the mean and the variance of normal distribution; ε is random numbers subject to standard normal distribution

The probability P_{2n} of the respondent n choosing a free road can be expressed as:

$$P_{2n} = \int \frac{e^{V_{2n}(\alpha,\beta,\chi)}}{\sum_{j=1}^{2} e^{V_{jn}(\alpha,\beta,\chi)}} f(\chi) d\chi \tag{7}$$

In the formula above, $f(\chi)$ is the density function of the parameter χ.

4.2 Random Numbers for Simulating the Parameter χ of Fatal Risk

500 random numbers obeying standard normal distribution are generated marked as $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{500}$. Thus, 500 random numbers of the parameters χ which obeys negative truncated normal distribution can be obtained marked as $\chi_1 = -\max(0, \mu + \sigma\varepsilon_1)$, $\chi_2 = -\max(0, \mu + \sigma\varepsilon_2), \ldots, \chi_{500} = -\max(\mu + \sigma\varepsilon_{500})$. Similarly, 500 random numbers of the parameters χ which obeys negative lognormal distribution can be obtained marked as $\chi_1 = -e^{\mu+\sigma\varepsilon_1}$, $\chi_2 = -e^{\mu+\sigma\varepsilon_2}, \ldots, \chi_{500} = -e^{\mu+\sigma\varepsilon_{500}}$.

4.3 Simulation Probability of Selection Probability

The simulation probability of the respondent n choosing a toll road can be expressed as:

$$\hat{P}_{1n} = \frac{1}{500} \sum_{i=1}^{500} \frac{e^{V_{1n}(\alpha,\beta,\chi_i)}}{\sum_{j=1}^{2} e^{V_{jn}(\alpha,\beta,\chi_i)}} \tag{8}$$

The simulation probability of the respondent n choosing a free road can be expressed as:

$$\hat{P}_{2n} = \frac{1}{500} \sum_{i=1}^{500} \frac{e^{V_{2n}(\alpha,\beta,\chi_i)}}{\sum_{j=1}^{2} e^{V_{jn}(\alpha,\beta,\chi_i)}} \tag{9}$$

4.4 Logarithmic Likelihood Function

$$L = \sum_{n=1}^{N} \left(\delta_{1n} \ln \hat{P}_{1n} + \delta_{2n} \ln \hat{P}_{2n} \right) \tag{10}$$

In the formula above: $\delta_{jn} = \begin{cases} 1, respondent\ n\ choosing\ alternative\ j; \\ 0,\ others \end{cases}$

Maximum likelihood estimation method is used to determine parameter estimates when the logarithmic likelihood function L is maximal. The accuracy of ML models can be determined by statistics of t-value, hit ratio and $\bar{\rho}^2$.

5 Data Acquisition and Parameter Calibrations

5.1 Data Acquisition

Private car commuters from Dalian to Lvshun were investigated. The survey sites included enterprises, schools and shopping malls and the survey way was interviewing investigators. From March to April in 2018, 500 questionnaires were sent out. Then 407 questionnaires were recovered, 324 of which were valid with the effective response rate close to 80%. According to foreign research experience, the ML model needs at least 300–500 sample values. In our survey, 972 (324 × 3) sample values were obtained, which meets the requirement of parameter calibrations.

Table 4 Parameter estimation results of model 1 and model 2

ML models Parameters	ML-1	ML-2
	Estimation values (t-value)	Estimation values (t-value)
Constant a_0	−2.2091(−1.963*)	−2.1856(−2.156*)
Travel cost coefficient α	−0.2543(−2.377*)	−1.3121(−3.322**)
Travel time coefficient β	−0.1724(−2.528*)	−0.6252(−2.892**)
Fatal risk coefficient μ	2.8112(2.496*)	2.2870(3.−377**)
Fatal risk coefficient σ	0.8512(2.259*)	0.6761(2.586**)
$L(\beta)$	−363.1923	−326.6894
$\bar{\rho}^2$	0.1516	0.2269
Hit ratio (%)	70.42	81.70

Note: **indicates that the significance test level is 1%, * indicates that the significance test level is 5%

5.2 Parameter Calibrations

According to the survey data, Monte Carlo simulation algorithm and Gauss 9.0 software were used to calibrate and verify the parameters. In order to eliminate the error of single simulation, 150 simulations were carried out and their average values were used as estimation values of the parameters. The calibration results of parameters are shown in Table 4.

5.3 Comparative Analysis Between ML Models

According to calibration results of model 1 showed in Table 4, the parameter α and β are both negative, which indicates that the higher travel cost or travel time is, the lower the utility function is. This conclusion is consistent with the hypothetical condition. As for t-values, all the absolute values are greater than 1.96. That is to say all parameters pass the significance test, which shows that the parameter settings are reasonable. However, the $\bar{\rho}^2$ is 0.1516 that does not meet the accuracy requirement in the area (0.2–0.4). Moreover, the hit ratio is 70.42% that does not meet the accuracy requirement of more than 80%. Therefore, the accuracy of ML model 1 with the fatal risk parameter obeying truncated normal distribution is still low, although it overcomes the shortcomings of normal distribution model.

According to calibration results of model 2 showed in Table 4, the parameter α and β are both negative, which indicates the parameter evaluation results are consistent with the Hypothetical condition. With regard to t-values, all the absolute values are greater than 1.96, which also shows that the parameter settings are reasonable. Moreover, the $\bar{\rho}^2$ is 0.2269 and the hit ratio is 81.70%, both of which meet the

Table 5 Utility functions of ML models with the cost parameter obeying lognormal distribution

ML models	Hypothetical condition	Fixed items of utility function of ML models
ML-3	① Fatal risk coefficient χ is a negative constant ② Travel time coefficient β is a negative constant ③ Travel cost coefficient α obeys negative lognormal distribution	$V_{1n} = a_0 - \exp(\mu + \varepsilon\sigma)c_{1n} + \beta t_{1n} + \chi r_{1n}$ $V_{2n} = \beta t_{2n} + \chi r_{2n}$
ML-4	① Fatal risk coefficient χ obeys negative lognormal distribution ② Travel time coefficient β is a negative constant ③ Travel cost coefficient α obeys negative lognormal distribution	$V_{1n} = a_0 - \exp(\mu_2 + \varepsilon_2\sigma_2)c_{1n} + \beta t_{1n} - \exp(\mu_1 + \varepsilon_1\sigma_1)r_{1n}$ $V_{2n} = \beta t_{2n} - \exp(\mu_1 + \varepsilon_1\sigma_1)r_{2n}$

Note μ and σ are respectively the mean and the variance of normal distribution; ε is random numbers subject to standard normal distribution

accuracy requirements. Therefore, model 2 with the fatal risk parameter obeying lognormal distribution is superior to model 1, and has a high accuracy.

6 Improved ML Models and Analysis of Calibration Results

6.1 Improved ML Models

Because VOSL is determined by coefficients of both fatal risk and travel cost, it is necessary to analyze the different influence of travel cost on different travelers. As 4.3 showed, it is most reasonable that the fatal risk coefficient obeys negative lognormal distribution, so the travel cost coefficient may also be assumed to obey negative lognormal distribution. Then two kinds of improved ML models (model 3 and model 4) are proposed, with their parameter assumptions and utility functions shown in Table 5.

6.2 Calibration Results of Improved ML Models

Similarly, the Monte Carlo simulation algorithm is used for 150 simulations, and the average value of simulation results is used as parameter estimation values (see Table 6).

Table 6 Parameter estimation results of model 3 and model 4

Variables	ML-3		ML-4	
	Parameters	Estimation values (t-value)	Parameters	Estimation values (t-value)
Constant	a_0	−2.2593(−2.063*)	a_0	−3.1568(−2.514*)
Travel cost	μ	0.1978(2.587**)	μ_2	0.2190(4.322**)
	σ	0.7041(3.142**)	σ_2	0.6015(3.341**)
Travel time	β	−0.4173(−3.528**)	β	−0.5158(−4.275**)
Fatal risk	χ	−8.7450(−3.236**)	μ_1	2.2812(3.017**)
			σ_1	0.3041(2.608**)
$L(\beta)$	−330.9108		−321.2443	
$\bar{\rho}^2$	0.2078		0.2534	
Hit ratio (%)	80.14		84.76	

Seen from calibration results of model 3 showed in Table 6, the parameter α and β are both negative, which is consistent with the hypothetical condition. T-values shows that all parameters pass the significance test. The $\bar{\rho}^2$ is 0.2078 and the hit ratio is 80.14%, which indicates that ML model 3 with the travel cost parameter obeying lognormal distribution has a high accuracy.

6.3 VOSL Distribution Function and Its Mathematical Expectation

Calibration results in Tables 4 and 6 show that model 2, model 3 and model 4 all have high accuracy. Therefore, distribution functions of VOSL are further deduced based on model 2, model 3 and model 4. The mathematical expectations (see Table 7) and the distribution regulations (shown in Figs. 1, 2 and 3) of VOSL are finally obtained. It is easy to prove that VOSL of model 2 obeys lognormal distribution with parameters $(\mu - \ln(-\alpha), \sigma^2)$, VOSL of model 3 obeys lognormal distribution with

Table 7 Distribution functions and mathematical expectations of VOSL

ML models	VOSL distribution function	Parameters of distribution function	Mathematical expectations(10^5 RMB)
ML-1	lognormal distribution	LOGNORM (2.0154, 0.6761^2)	9.4306
ML-2	lognormal distribution	LOGNORM (1.9707, 0.7047^2)	9.1941
ML-3	lognormal distribution	LOGNORM (2.0622, 0.6740^2)	9.8684

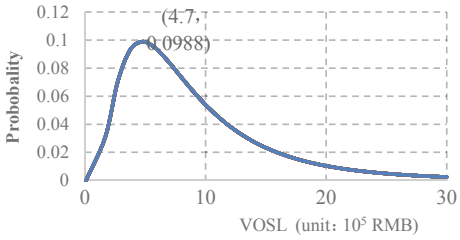


Fig. 1 VOSL distribution of model 2

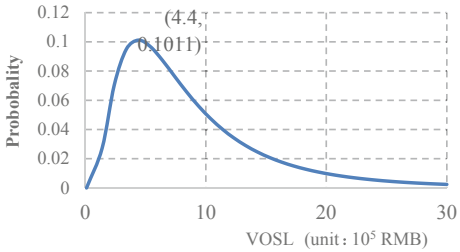


Fig. 2 VOSL distribution of model 3

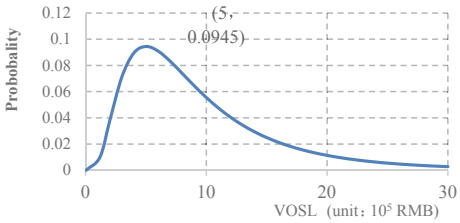


Fig. 3 VOSL distribution of model 4

parameters $(\ln(-\chi) - \mu, \sigma^2)$, and VOSL of model 4 obeys lognormal distribution with parameters $(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

As seen from Table 7 and Fig. 1, the mathematical expectation of VOSL in model 2 is $\exp(2.0154 + 0.6761^2/2) = 9.4306$, that is 943,060 RMB. The probability increases first and then decreases with the increase of VOSL estimates, and reaches the maximum (9.88%) when the estimate is 470,000 RMB.

As seen from Table 7 and Fig. 2, the mathematical expectation of VOSL in model 3 is $\exp(1.9707 + 0.7047^2/2) = 9.1941$, that is 919,410 RMB. The probability increases first and then decreases with the increase of VOSL estimates, and reaches the maximum (10.11%) when the estimate is 440,000 RMB.

As seen from Table 7 and Fig. 3, the mathematical expectation of VOSL in model 4 is $\exp(2.0622 + 0.6740^2/2) = 9.8684$, that is 986,840 RMB. The probability increases first and then decreases with the increase of VOSL estimates, and reaches the maximum (9.45%) when the estimate is 500,000 RMB.

6.4 Comparative Analysis of ML Models and Evaluation Results

Model 2, model 3 and model 4 were comparatively analyzed from calibration results of Tables 4 and 6. ML Model 4 has the largest logarithmic likelihood function (−321.2443), the largest $\bar{\rho}^2$(0.2534) and the largest hit ratio (84.76%). Therefore, model 4 is superior to model 2 and model 3. That is to say, ML model 4 with coefficients of fatal risk and travel cost simultaneously obeying lognormal distribution has the highest accuracy. ML Model 4 not only analyzes the differential impact of fatal risk, but also that of travel cost. Therefore, ML model 4 can explain the distribution characteristics of VOSL more reasonably.

The VOSL estimate based on model 4 is 986,840 RMB. Compared with the VOSL estimate (513,000 RMB) in Beijing assessed by CVM in 2008 [16], our estimate is higher, which is mainly due to different evaluation methods and longer time interval between data acquisition. At present, many scholars have made comprehensive analysis of VOSL samples with different time points and different evaluation methods, pointing out that "VOSL is affected by personal wealth, evaluation methods and other factors" [17, 18] Compared with the VOSL estimate (3.72/3.28 million RMB) in Nanjing assessed by SCM and ML models in 2016 [14], our estimate is lower, which is mainly because of economic and cultural differences among regions besides the difference of parameter distribution function. Through the comparative study of VOSL among different regions, many scholars draw the conclusion that the higher the level of regional economy and culture is, the higher the VOSL estimate is [19].

In order to further verify the model reliability, value of travel time (VTT) was evaluated. According to the formula $VTT = \beta/\alpha$, it is easy to prove VTT based on model 4 obeys lognormal distribution with parameters $(\ln(-\beta) - \mu_2, \sigma_2^2)$, marked as VTT ~ LOGNORM $(-0.8810, 0.6015^2)$. The mathematical expectation of VTT is EXP $(-0.8810 + 0.6015^2/2) = 0.4965$ (RMB/minute), which is 29.79 RMB per hour. It is basically consistent with the results of many related studies, where most VTT estimates are between 20 and 40 RMB per hour [20, 21].

7 Conclusion

Firstly, VOSL evaluation model based on 4 ML models was constructed, taking truncated normal and lognormal distribution as parameter distribution functions. Then SCM questionnaire was designed to conduct traffic survey in Dalian, and Monte Carlo method was used to calibrate and compare ML models. Finally, the VOSL estimate and its distribution curve are obtained. There are two main conclusions: (1) ML model 4 with coefficients of fatal risk and travel cost simultaneously obeying lognormal distribution has the highest accuracy, which is more reasonable to explain the changing regulation of VOSL. (2) VOSL based on ML model 4 obeys lognormal distribution marked as VOSL ~ LOGNORM $(2.0622, 0.6740^2)$. The mathematical

expectation of VOSL is 986,840 RMB, and its maximum probability (9.45%) is reached when it is 500,000 RMB. The research deficiency is that lognormal distribution has unilateral constraints, which causes the defect of too long tail. The future research direction is to construct ML models whose parameters obey S_B distribution with bilateral constraints.

Foundations National Natural Science Foundation of China (51608088); The Youth Program of Humanities and Social Sciences of the Ministry of Education (16YJC630075); Social Science Planning Fund of Liaoning Province (L15BGL003); Doctoral Research Initiation Fund of Liaoning Province (201601261). Author resume: LIU Wen-ge(1981-), female, Dalian Liaoning Province, associate professor, PhD, Email: 18737653@qq.com.

References

1. Schelling TC (1968) The life you save may be your own. In: Problems in public expenditure analysis. Brookings, Washington, DC, pp 127–162
2. Xing YE, Le-jia DU (2017) On the measurement of the value of life. J Ningbo Univ (Liberal Arts Edition) 30(5):117–122
3. Carson RT, Flores NE, Meade NF (2001) Contingent valuation: controversies and evidence. Environ Resour Econ 19:173–210
4. Shen JY (2005) A review of stated choice method[D]. Osaka University, Osaka
5. Rizzi LI, Ortúzar J (2003) Stated preference in the valuation of interurban road safety. Accid Anal Prev 35(1):9–22
6. Iragüen P, Ortúzar J (2004) Willingness-to-pay for reducing fatal accident risk in urban areas: an internet-based web page stated preference survey. Accid Anal Prev 36(1):513–524
7. Zhao S, Zhang Y (2008) Disaggregate model based life value statistics analysis. Technol Highway Transp 5:132–135
8. Liu W, Zhao S (2013) The value of statistical life in road traffic based on logit model. J Transp Syst Eng Inf Technol 13(1):137–141
9. Guan H (2004) Disaggregate model—a tool of traffic behavior analysis. China Communication Press, Beijing
10. McFadden D, Train K (2000) Mixed MNL models for discrete response. J Appl Econometrics 15:447–470
11. Rouwendal J, de Blaeij A, Rietveld P, Verhoef E (2009) The information content of a stated choice experiment: a new method and its application to the value of a statistical life. Transp Res Part B 4:1–16
12. Rheinberger CM (2011) A mixed logit approach to study preferences for safety on alpine roads. Environ Resource Econ 49(1):121–146
13. Liu W, Zhao S (2013) The evaluation study on the value of a statistical life in road traffic based on mixed logit model. Stat Inf Forum 28(2):109–112
14. Yang Z, Liu P, Xu X (2016) Estimation of social value of statistical life using willingness-to-pay method in Nanjing, China. Accid Anal Prev 95:308–316
15. Cha J, Cho BR, Sharp JL (2013) Rethinking the truncated normal distribution. J Exp Des Process Optim 3(4):327–363
16. Jun-peng LUO, Yong HE (2008) Application of contingent valuation method in evaluation of value of statistical life in road safety. J Highway Transp Res Dev 6:134–138
17. De Blaeij A, Florax R, Rietveld P, Verhoef E (2003) The value of statistical life in road safety: a meta-analysis. Accid Anal Prev 35(6):973–986

18. Hoffmann S, Krupnick A, Qin P (2017) Building a set of internationally comparable value of statistical life studies: estimates of Chinese willingness to pay to reduce mortality risk. J Benefit-Cost Anal 8(2):251–289
19. Zan H, Scharff RL (2017) Regional differences in the value of statistical life. J Consum Policy 40:157–176
20. Zhao S, Wang X, Yao R, Zhang Y (2010) Value of driver commuting travel time based on mixed logit model. J Jilin Univ (Engineering and Technology Edition) 40(2):406–411
21. Zhou Y (2016) The value of time calibration and its application in traffic distribution prediction-taking Dongguan as an example. Shenzhen University, Shenzhen

Research on Point-to-Point Direct Transportation Organization Mode of Railway Bulk Goods

Wei Lu

Abstract This paper firstly gives the definition of point-to-point direct transportation products for railway bulk cargoes, and then analyzes the organization conditions of point-to-point direct transportation. Finally, the total vehicle hour consumption of the loading and unloading traffic in the whole transportation process is taken as the objective function to found Point-to-point direct traffic flow organization model, and corresponding cases are given for analysis to verify the feasibility of the model.

Keywords Railway freight · Point-to-point direct · Car flow organization

1 Introduction

With the continuous expansion of the current railway network, railway transportation will be further developed. Considering the advantages of large volume, low cost and accurate time for railway cargo transportation, for the transportation of bulk cargo such as coal, oil and grain, rail transport will be the preferred mode of transport. Moreover, for bulk cargoes, due to their large volume of transportation, inventory time and cycle requirements, when carrying out railway transportation, the organization of relevant traffic flow should be carried out as reasonably as possible, and a reasonable traffic organization scheme should be established to ensure its reasonable and efficient transportation.

For the railway traffic organization, experts and scholars at home and abroad have carried out a lot of analysis and research on this, and have achieved fruitful results. In the existing research field, Porovoy [1] is the earliest use of systematic thinking, the "simultaneous calculation method" to optimize the technical station traffic flow scheme research. It established the model using technical station traffic flow adaptation information and whether the loading traffic is organized directly as decision

W. Lu (✉)
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
e-mail: 17120853@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_69

variables. Lin et al. [2] combines the characteristics of actual traffic flow organization at the present stage to analyze the necessary conditions and parameter selection of the organization's originating direct trains, and constructs a nonlinear 0–1 model of the train formation scheme of the loading place. On this basis, the comprehensive optimization model of multiple loading trains, the simultaneous optimization model of the train formation plan and the technical plan of the train station, and the multi-objective optimization of the heavy-duty direct train plan are also constructed. It comprehensively analyzes the theoretical problems of many aspects of the train formation plan of the loading place. In [3, 4], the transportation and inventory costs are included in the same objective function. The minimum cost of the logistics system is used to construct the direct-to-cargo train operation plan. The nonlinear 0–1 planning model is used to determine the train operation of the loading place. best plan. Lin [5] analyzed the composition of the staying time of the vehicle through the technical station, and converted the operating cost, the non-adjusted operating cost and the waiting time of the vehicle into a unified unit to construct the objective function. Cao et al. [4] and Ji et al. [6] considered logistics costs and inventory costs, respectively. In [7], the possibility and rationality of direct trains in the organization of loading places are analyzed, and on the basis of the selection of cost parameters, a nonlinear 0-1 planning model for the direct train plan of the loading area is constructed. The model fully considers the various combinations of traffic flow in the loading place, and incorporates the transportation and inventory costs into the same objective function, and determines the freight train operation plan with the optimal logistics system cost. In [8], from the perspective of network flow and combination optimization, based on the OD traffic flow between stations and stations, the relationship between the initial traffic flow and the technical traffic flow is analyzed, and the hypothesis of existing trains between adjacent fulcrums are abandoned. It considers the comprehensive optimization problem of the direct train formation plan, the technical direct formation plan and the section train operation plan of the road network loading place. Fan et al. [9] analyzed the types of direct trains originating from railway freight, the favorable indicators and the conditions of the driving and the determination of the plan, etc., and constructed the mathematical model of the initial direct transportation grouping plan.

2 Problem Description

(1) **Connotation of direct transportation of railway bulk goods from point to point**

① One or several loading points at the same station, the direct trains consisting of vehicles loaded by one or several delivery units;
② Unload the vehicle at one or several unloading locations at the same unloading station.

In addition, the direct-to-point train loading location can be a direct circulation of the fixed vehicle bottom, or a non-circulating direct transmission without fixing the vehicle bottom; It can be a direct or fixed weight direct change of weight; It can be a direct train that runs regularly every day or a direct train that runs irregularly.

(2) **Point-to-point direct transport conditions**

Organizing point-to-point direct transportation can speed up the transportation of materials and the turnover of rolling stock, so that the technical and economic effects of expanding transportation capacity and reducing transportation costs can be obtained.

Organizing point-to-point direct trains is not economically viable under any conditions. For example, in the case of insufficient storage capacity, insufficient daily production or consumption, poor handling equipment, and long time for loading and unloading, the organization of direct trains may result in a large backlog of materials and extended truck stays. It will even bring great difficulties to enterprises and loading and unloading stations.

The organization of point-to-point direct trains should meet the following basic conditions:

① The station (referring to a dedicated line or a number of loading points) must have a certain amount of freight and direct traffic, and the capacity of the library, the loading capacity and the length of the wiring can guarantee the whole listed car.
② The yard capacity and unloading capacity of the unloading station can guarantee the requirements of the entire column to the unloading;
③ Single-vehicle flow or combined traffic flow-to-point direct trains, the total loading and unloading waiting hour loss of the direct-moving traffic flow should generally be less than the consumption according to the technical traffic flow organization;
④ For the direct trains whose organization needs to change the weight standard during the operation, there should be equipment conditions for the technical work of picking up the trailers at the weight-changing station, and suitable traffic for the axles of the under-axis train.

3 Establish Model

(1) **The parameters introduced are as follows**:

N_{st}: Daily average planned traffic from the loading location s to the unloading site t, unit: car;

m_{st}: The maximum number of dense loading and unloading trucks allowed for N_{st}, unit: car;

$\omega_s^{direct\_ori}$: The starting vehicle flow is consumed in the loading area by the vehicle leaving the loading area in a direct manner;

$\omega_t^{direct\_\mathrm{des}}$: The hourly consumption of each vehicle entering the unloading area from the direct flow of the vehicle;

$\omega_s^{local\_ori}$: The initial traffic flow is consumed per vehicle that leaves the loading area in a non-direct manner;

$\omega_t^{local\_\mathrm{des}}$: The hourly consumption of the vehicle is not indirect mode and enters the unloading area;

τ_i^{change}: The unit cost of the initial traffic flow at the station i or weight change

τ_i^{sort}: The unit cost of the initial non-direct traffic flow at the station i;

V: The collection of technical stations on the road network;

$V(s)$: The collection of technical stations adjacent to the loading site;

$V(s,t)$: The traffic flow N_{st} is on a prescribed path, and the technical station passes along the way (excluding the technical station adjacent to s);

$\rho(s,t)$: The collection of technical stations through which traffic n may pass;

$V(k,t)$: The collection of technical stations that may be adapted after the adaptation of the traffic flow N_{st} at the station k, $k \in V(k,t)$;

$Q(s)$: All the cars that are sent from the loading place s are a collection of unloading places.

(2) **Define the following decision variables**:

$$x_{st} = \begin{cases} 1 \text{ If you drive a cargo train that goes directly from the loading place to the unloading place} \\ 0 \text{ other} \end{cases}$$

$$y_{st}^k = \begin{cases} 1 \text{ The technical station that the car flow } N_{st} \text{ was first adapted is } k \\ 0 \text{ other} \end{cases}$$

(3) **Establish model**:

The converted car hourly consumption for different traffic organization schemes at loading place is:

① When the traffic flow N_{st} is a large-scale traffic flow, and the direct train is directly driven to the unloading place, the converted car hourly consumption at the loading place and the unloading place and during the transfer operation is:

$$Z^{node-node} = N_{st}\left(\omega_s^{direct\_ori} + \omega_t^{direct\_\mathrm{des}} + \sum_{i\in\rho(s,t)} \delta_{st}^i \tau_i^{change}\right) x_{st} \tag{1}$$

Among them

$$\delta_{st}^i = \begin{cases} 1, \text{ If the traffic flow } N_{st} \text{ is performing the weight change operation at the station } i \\ 0, \text{ If the traffic flow } N_{st} \text{ is not carried out at the station } i \end{cases}$$

Whether to carry out the weight change operation is determined by factors such as the number of traction of different traction sections. It is now stipulated that each technical station passing through the direct train has a range of the number of fixed allowed traffic. If the traffic flow N_{st} is not within the interval, the weight change operation is required at the technical station, otherwise the weight change operation is not required, thereby satisfying the traction limit of the different traction sections.

② When the traffic flow N_{st} does not organize the direct train to the unloading place, in addition to the corresponding parking hour consumption at the loading and unloading sites, it also needs to be adapted at certain technical stations along the way. At least in the last technical station, the station performs the disassembly work and consumes the corresponding car hours. Assuming that the traffic path has been given and the first reconfiguration station is station k, the technical station that needs to be adapted along the way is determined under the condition that the technical station train formation plan is known. The set of technical stations that need to be adapted is $V(k, t)$, and the k-station is included in $V(k, t)$. If the traffic flow N_{st} does not organize the direct train, the converted car hourly consumption is:

$$Z^{sort} = N_{st} \cdot \sum_{k \in \rho(s,t)} y_{st}^{k} \left(\omega_{s}^{local\_ori} + \omega_{t}^{local\_\mathrm{des}} + \sum_{k' \in V(k,t)} \tau_{k'}^{sort} \right) \tag{2}$$

Thus, the total consumption of traffic N_{st} on the way is:

$$\begin{aligned} Z_{st} = {} & N_{st} \left(\omega_{s}^{direct\_ori} + \omega_{t}^{direct\_\mathrm{des}} + \sum_{i \in \rho(s,t)} \delta_{st}^{i} \tau_{i}^{change} \right) x_{st} \\ & + N_{st} \cdot \sum_{k \in \rho(s,t)} y_{st}^{k} \left(\omega_{s}^{local\_ori} + \omega_{t}^{local\_\mathrm{des}} + \sum_{k' \in V(k,t)} \tau_{k'}^{sort} \right) \end{aligned} \tag{3}$$

For traffic flow N_{st}, there are 2 ways to get it to your destination. Either organize a direct train to the unloading site; or send it to a technical station to be incorporated into the corresponding technical traffic. Therefore, the unique conditions for the traffic flow N_{st} organization scheme are:

$$x_{st} + \sum_{k \in \rho(s,t)} y_{st}^{k} = 1 \tag{4}$$

Introduced to allow the number of dense loading and unloading trucks m_{st}, and the following conditions are met:

$$m_{st} = \min\left\{m_{st}^{u}, m_{st}^{d}\right\} \tag{5}$$

Among them, m_{st}^{u} and m_{st}^{d} respectively indicates the number of intensive loading and unloading.

Dense loading and unloading refers to the continuous feeding or unloading process, which is neither stopped due to insufficient cargo volume nor stopped due to insufficient capacity of the yard storage. It is not only related to the capacity of the cargo space, but also related to the merger of traffic flow.

When the loading place s is programmed to the starting train of a fulcrum station k (excluding the technical station adjacent to the loading place), the total number of permitted dense loading and unloading vehicles to the attracted vehicles is met the following requirements:

$$\overline{m_{sk}} I\left(\sum_{t\in Q(s)} y_{st}^{k}\right) - \sum_{t\in Q(s)} m_{st} y_{st}^{k} \le 0 \quad k \in V(s,t) \tag{6}$$

where $I(x)$ is a step function, defined as

$$I(x) = \begin{cases} 1 & x > 0 \\ 0 & x \le 0 \end{cases} \tag{7}$$

In this way, the total vehicle hour consumption in the entire transportation is the objective function and minimized. Get the mathematical model of the following train-to-drive train plan:

$$\begin{aligned} &\text{(M-I)} \min \sum_{t\in Q(s)} Z_{st} \\ &\text{S. T.} \\ &\quad x_{st} + \sum_{k\in\rho(s,t)} y_{st}^{k} = 1 && t \in Q(s) \\ &\quad \overline{m_{sk}} I\left(\sum_{t\in Q(s)} y_{st}^{k}\right) - \sum_{t\in Q(s)} m_{st} y_{st}^{k} \le 0 && k \in V(s,t) \\ &\quad x_{st}, y_{st}^{k} \in \{0,1\} && \forall t, k \end{aligned}$$

The above constraint (6) is now processed to be converted into a linear constraint. Introducing variable

$$y_{sk} = I\left(\sum_{t\in Q(s)} y_{st}^{k}\right) = \begin{cases} 1 & \text{If there is a point-to-point direct train from the loading place to the technical station } k \\ 0 & \text{otherwise} \end{cases}$$

$k \in V(s,t)$. Bringing y_{sk} into the above constraint (6), and performing the relevant transformation can get the following constraints:

$$\overline{m_{sk}} y_{sk} - \sum_{t \in Q(s)} m_{st} y_{st}^{k} \leq 0 \quad k \in V(s,t) \tag{8}$$

$$M y_{sk} - \sum_{t \in Q(s)} y_{st}^{k} \geq 0 \quad k \in V(s,t) \tag{9}$$

$$y_{sk} \in \{0,1\} \quad k \in V(s,t) \tag{10}$$

M in the above formula satisfies the condition $M \geq n$, and n represents the number of elements in the $Q(s)$.

Therefore the model (M-I) can be transformed into the following linear 0-1 planning form:

$$\text{(M-II)} \quad \min \sum_{t \in Q(s)} Z_{st}$$

S. T.

$$\begin{aligned}
&x_{st} + \sum_{k \in \rho(s,t)} y_{st}^{k} = 1 && t \in Q(s) \\
&\overline{m_{sk}} y_{sk} - \sum_{t \in Q(s)} m_{st} y_{st}^{k} \leq 0 && k \in V(s,t) \\
&y_{sk} \in \{0,1\} && k \in V(s,t) \\
&x_{st}, y_{st}^{k} \in \{0,1\} && \forall t, k
\end{aligned}$$

4 Case Analysis

A simple road network structure diagram is given here, as shown in Fig. 1. 1 means

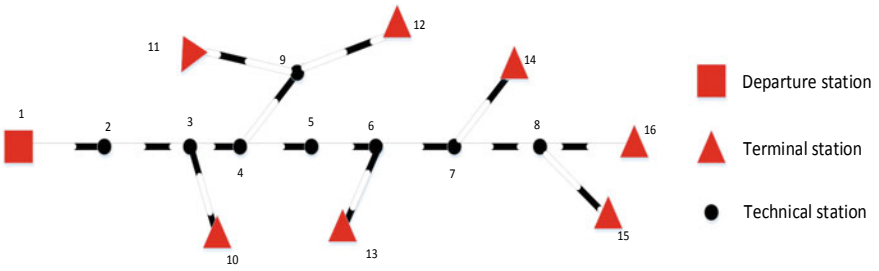


Fig. 1 Road network structure

the station at the originating station, 2–9 means the technical station on the way, and 10–14 means the station is unloaded at the end.

The relevant data and parameters are shown in Table 1.

The technical parameters of each technical station are shown in Table 2.

It is stipulated that if the point-to-point direct traffic N_{st} is within the allowed traffic flow range of each technical station, no weight change operation is required; if the traffic flow N_{st} is smaller than the lower limit of the technical station traffic range, the weight change operation is increased to the lower limit traffic flow of the allowed traffic flow range; if the traffic flow N_{st} is greater than the upper limit of the technical station traffic range, it needs to be reduced to the upper limit traffic flow of the allowed traffic flow range.

In addition, the remaining parameters are specified as follows: when organizing point-to-point direct and adapted delivery, the unit hour consumption at the loading and unloading site is $\omega_s^{direct\_ori} = 8$, $\omega_t^{direct\_\mathrm{des}} = 8$, $\omega_s^{local\_ori} = 6$, $\omega_t^{local\_\mathrm{des}} = 6$.

According to the grouping plan of the technical station on the way, it is possible to determine the set of technical stations $V(k, t)$ to be adapted along the way of each vehicle flow, as follows:

Table 1 Traffic flow OD data

Traffic flow N_{st}	Number of driving (car/day)	m_{st} (car)
$N_{1,10}$	12	50
$N_{1,11}$	20	50
$N_{1,12}$	50	50
$N_{1,13}$	45	50
$N_{1,14}$	36	50
$N_{1,15}$	66	50
$N_{1,16}$	55	50

Table 2 Related parameters of the technical station

Technical station number	2	3	4	5	6	7	8	9
Technical station adaptation operation consumption τ_i^{sort}	5	3	4	3	4	4	5	5
Technical station weight change operation τ_i^{change}	1	1.5	1	2	1.5	1	1.5	1
Technical station limited by the number of traction limits	30–60	35–45	40–50	45–50	50–60	35–45	50V55	30–40
$\overline{m_{sk}}$	40	40	40	40	40	40	40	40

$N_{1,10}: V(2,10) = \{2,3\}, V(3,10) = \{3\};$
$N_{1,11}: V(2,11) = \{2,3,9\}, V(3,11) = \{3,4,9\}, V(4,11) = \{4,9\}, V(9,11) = \{9\};$
$N_{1,12}: V(2,12) = \{2,4,9\}, V(3,12) = \{3,9\}, V(4,12) = \{4,9\}, V(9,12) = \{9\};$
$N_{1,13}: V(2,13) = \{2,4,5\}, V(3,13) = \{3,5,6\},$
$V(4,13) = \{4,6\}, V(5,13) = \{5,6\}, V(6,13) = \{6\};$
$N_{1,14}: V(2,14) = \{2,3,5,6\}, V(3,14) = \{3,5,6\}, V(4,14) = \{4,6,7\},$
$V(5,14) = \{5,7\}, V(6,14) = \{6,7\}, V(7,14) = \{7\};$
$N_{1,15}: V(2,15) = \{2,3,5,6,8\}, V(3,15) = \{3,4,6,8\}, V(4,15) = \{4,6,7,8\},$
$V(5,15) = \{5,7,8\}, V(6,15) = \{6,7\}, V(7,15) = \{7,8\}, V(8,15) = \{8\};$
$N_{1,16}: V(2,16) = \{2,4,5,6,7\}, V(3,16) = \{3,5,6,7\}, V(4,16) = \{4,5,6,8\},$
$V(5,16) = \{5,6,8\}, V(6,16) = \{6,7\}, V(7,16) = \{7\}, V(8,16) = \{8\}.$

Then, the range of traffic allowed by the technical station limited by the traction number determines the set of technical stations that need to carry out the weight change operation on the way to each point-to-point traffic flow, as follows:

$$
\begin{aligned}
&N_{1,10}: \{2,3\};\\
&N_{1,11}: \{2,3,4\};\\
&N_{1,12}: \{3,9\};\\
&N_{1,13}: \{6\};\\
&N_{1,14}: \{4,5,6,7\};\\
&N_{1,15}: \{2,3,6,7,8\};\\
&N_{1,16}: \{3,6,7,8\}.
\end{aligned}
$$

The point-to-point traffic organization model for this case is listed below based on the previous section:

$$Z = \min\left(Z_{1,10} + Z_{1,11} + Z_{1,12} + Z_{1,13} + Z_{1,14} + Z_{1,15} + Z_{1,16}\right)$$

S. T.

$$
\begin{aligned}
x_{1,10} + y_{1,10}^{2} + y_{1,10}^{3} &= 1\\
x_{1,11} + y_{1,11}^{2} + y_{1,11}^{3} + y_{1,11}^{4} + y_{1,11}^{9} &= 1\\
x_{1,12} + y_{1,12}^{2} + y_{1,12}^{3} + y_{1,12}^{4} + y_{1,12}^{9} &= 1\\
x_{1,13} + y_{1,13}^{2} + y_{1,13}^{3} + y_{1,13}^{4} + y_{1,13}^{5} + y_{1,13}^{6} &= 1\\
x_{1,14} + y_{1,14}^{2} + y_{1,14}^{3} + y_{1,14}^{4} + y_{1,14}^{5} + y_{1,14}^{6} + y_{1,14}^{7} &= 1\\
x_{1,15} + y_{1,15}^{2} + y_{1,15}^{3} + y_{1,15}^{4} + y_{1,15}^{5} + y_{1,15}^{6} + y_{1,15}^{7} + y_{1,15}^{8} &= 1\\
x_{1,16} + y_{1,16}^{2} + y_{1,16}^{3} + y_{1,16}^{4} + y_{1,16}^{5} + y_{1,16}^{6} + y_{1,16}^{7} + y_{1,16}^{8} &= 1
\end{aligned}
$$

$$\bar{m}_{1,3}\, y_{1,3} - \sum_{t=10}^{16} m_{1,t}\, y_{1,t}^{3} \le 0$$

$$\bar{m}_{1,4}\, y_{1,4} - \sum_{t=11}^{16} m_{1,t}\, y_{1,t}^{4} \le 0$$

$$\bar{m}_{1,5}\, y_{1,5} - \sum_{t=13}^{16} m_{1,t}\, y_{1,t}^{5} \le 0$$

$$\bar{m}_{1,6}\, y_{1,6} - \sum_{t=13}^{16} m_{1,t}\, y_{1,t}^{6} \le 0$$

$$\bar{m}_{1,7}\, y_{1,7} - \sum_{t=14}^{16} m_{1,t}\, y_{1,t}^{7} \le 0$$

$$\bar{m}_{1,8}\, y_{1,8} - \sum_{t=15}^{16} m_{1,t}\, y_{1,t}^{8} \le 0$$

$$\bar{m}_{1,9}\, y_{1,9} - \sum_{t=11}^{12} m_{1,t}\, y_{1,t}^{9} \le 0$$

$$7\, y_{1,3} - \sum_{t=10}^{16} y_{1,t}^{3} \ge 0$$

$$6\, y_{1,4} - \sum_{t=11}^{16} y_{1,t}^{4} \ge 0$$

$$4\, y_{1,5} - \sum_{t=13}^{16} y_{1,t}^{5} \ge 0$$

$$4\, y_{1,6} - \sum_{t=13}^{16} y_{1,t}^{6} \ge 0$$

$$3\, y_{1,7} - \sum_{t=14}^{16} y_{1,t}^{7} \ge 0$$

$$2\, y_{1,8} - \sum_{t=15}^{16} y_{1,t}^{8} \ge 0$$

$$2\, y_{1,9} - \sum_{t=11}^{12} y_{1,t}^{9} \ge 0$$

$$y_{sk} \in \{0,1\} \quad k \in V(s,t)$$

$$x_{st}, \quad y_{st}^{k} \in \{0,1\} \quad k \in \rho(s,t)$$

The $Z_{1,10}$, $Z_{1,11}$, $Z_{1,12}$, $Z_{1,13}$, $Z_{1,14}$, $Z_{1,15}$, $Z_{1,16}$ in the objective function can be expressed as the following:

$$Z_{1,10} = N_{1,10}\left(\omega_s^{direct\_ori} + \omega_t^{direct\_\mathrm{des}} + \sum_{i\in\rho(1,10)} \delta_{1,10}^{i}\tau_i^{change}\right) x_{1,10} + N_{1,10}\cdot\sum_{k=2}^{3} y_{1,10}^{k}\left(\omega_s^{local\_ori} + \omega_t^{local\_\mathrm{des}} + \sum_{k'\in V(k,10)} \tau_{k'}^{sort}\right)$$

$$Z_{1,11} = N_{1,11}\left(\omega_s^{direct\_ori} + \omega_t^{direct\_\mathrm{des}} + \sum_{i\in\rho(1,11)} \delta_{1,11}^{i}\tau_i^{change}\right) x_{1,11} + N_{1,11}\cdot\sum_{k=2,3,4,9} y_{1,11}^{k}\left(\omega_s^{local\_ori} + \omega_t^{local\_\mathrm{des}} + \sum_{k'\in V(k,11)} \tau_{k'}^{sort}\right)$$

$$Z_{1,12} = N_{1,12}\left(\omega_s^{direct\_ori} + \omega_t^{direct\_\mathrm{des}} + \sum_{i\in\rho(1,12)} \delta_{1,12}^{i}\tau_i^{change}\right) x_{1,12} + N_{1,12}\cdot\sum_{k=2,3,4,9} y_{1,12}^{k}\left(\omega_s^{local\_ori} + \omega_t^{local\_\mathrm{des}} + \sum_{k'\in V(k,12)} \tau_{k'}^{sort}\right)$$

$$Z_{1,13} = N_{1,13}\left(\omega_s^{direct\_ori} + \omega_t^{direct\_\mathrm{des}} + \sum_{i\in\rho(1,13)} \delta_{1,13}^{i}\tau_i^{change}\right) x_{1,13} + N_{1,13}\cdot\sum_{k=2}^{6} y_{1,13}^{k}\left(\omega_s^{local\_ori} + \omega_t^{local\_\mathrm{des}} + \sum_{k'\in V(k,13)} \tau_{k'}^{sort}\right)$$

$$Z_{1,14} = N_{1,14}\left(\omega_s^{direct\_ori} + \omega_t^{direct\_\mathrm{des}} + \sum_{i\in\rho(1,14)} \delta_{1,14}^{i}\tau_i^{change}\right) x_{1,14} + N_{1,14}\cdot\sum_{k=2}^{7} y_{1,14}^{k}\left(\omega_s^{local\_ori} + \omega_t^{local\_\mathrm{des}} + \sum_{k'\in V(k,14)} \tau_{k'}^{sort}\right)$$

$$Z_{1,15} = N_{1,15}\left(\omega_s^{direct\_ori} + \omega_t^{direct\_\mathrm{des}} + \sum_{i\in\rho(1,15)} \delta_{1,15}^{i}\tau_i^{change}\right) x_{1,15} + N_{1,15}\cdot\sum_{k=2}^{8} y_{1,15}^{k}\left(\omega_s^{local\_ori} + \omega_t^{local\_\mathrm{des}} + \sum_{k'\in V(k,15)} \tau_{k'}^{sort}\right)$$

$$Z_{1,16} = N_{1,16}\left(\omega_s^{direct\_ori} + \omega_t^{direct\_\mathrm{des}} + \sum_{i\in\rho(1,16)} \delta_{1,16}^{i}\tau_i^{change}\right) x_{1,16}$$

$$+N_{1,16}\cdot\sum_{k=2}^{8}y_{1,16}^{k}\left(\omega_{s}^{local\_ori}+\omega_{t}^{local\_\mathrm{des}}+\sum_{k'\in V(k,16)}\tau_{k'}^{sort}\right)$$

Using the Lingo commercial software to solve the above model, the results are as follows: $x_{1,12}=1$, $x_{1,13}=1$, $y_{1,10}^{3}=1$, $y_{1,11}^{9}=1$, $y_{1,14}^{7}=1$, $y_{1,15}^{8}=1$, $y_{1,16}^{7}=1$. The rest of the variables have a value of 0. The cost of the solution is 4584.5 car hours. The corresponding traffic organization scheme is shown in Fig. 2.

In the above picture, the trains that arrive at the No. 12 and No. 13 unloading stations will be re-routed and delivered to other unloading areas. Among them, the train that was rerouted at No. 3 technical station on the way to the No. 10 unloading station; the train that was rerouted at No. 9 technical station on the way to the No. 11 unloading station; the train that was rerouted at No. 7 technical station on the way to

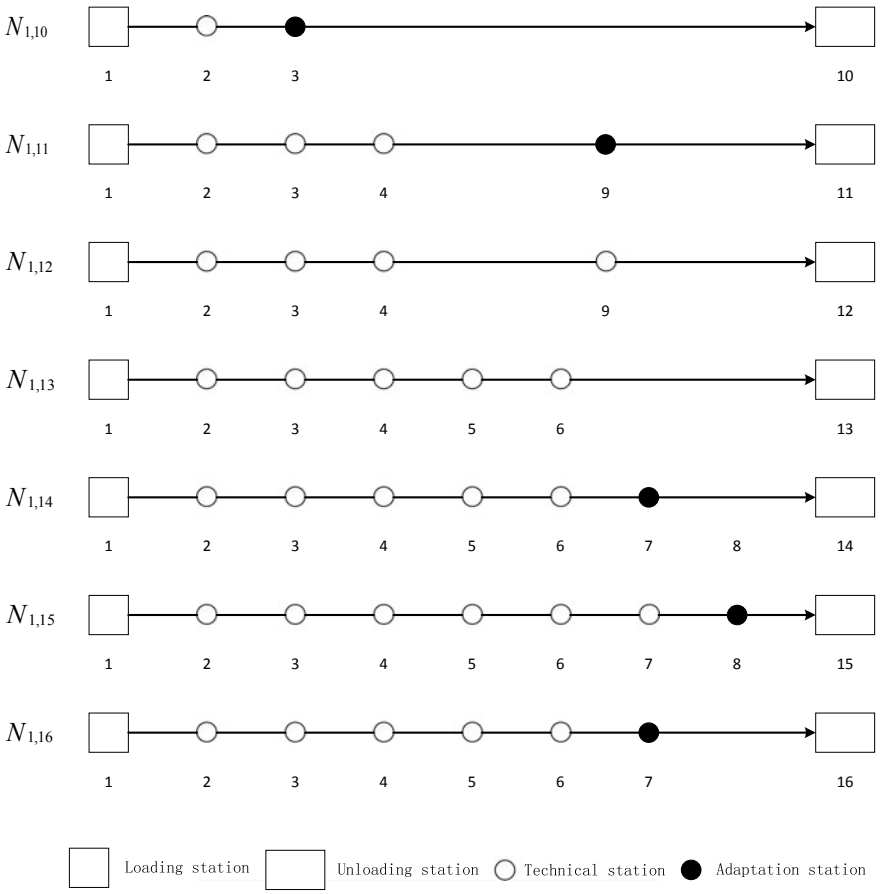


Fig. 2 Traffic organization plan

the No. 14 unloading station; the train that was rerouted at No. 8 technical station on the way to the No. 15 unloading station; the train that was rerouted at No. 7 technical station on the way to the No. 16 unloading station.

5 Conclusions

For the transportation of bulk cargo in railways, the form of traffic organization is different, which will result in different transportation and loading and unloading operating costs. Therefore, based on the point-to-point direct transportation organization, this paper analyzes the conditions of the point-to-point direct train, and then establishes the optimized organization model of the point-to-point direct train, and selects the point-to-point direct train according to its different restrictions. Other forms of trains. Finally, the relevant examples are given for analysis to verify the feasibility of the model.

References

1. Porovoy HE (1981) Direct transportation of goods. Teaching and Research of Driving Organization of Beijing Jiaotong University Room Translation. China Railway Publishing House, Beijing
2. Lin B, Zhu S, Shi D, He S (1995) Optimization model of direct train formation plan in loading area. China Railway Sci (2):108–114
3. Cao X, Lin B, Liu H, Yan H (2007) Optimization of vehicle flow organization in direct base. J China Railway Soc (1):16–20
4. Cao X, Lin B, Yan H (2006) Optimization model of direct train operation scheme in loading area. J China Railway Soc (4):6–11
5. Lin B (1996) Optimization of train formation plan with nonlinear adaptation costs. J China Railway Soc S1:37–42
6. Ji L, Lin B, Wang Z (2009) Study on optimization model of loading and unloading traffic flow based on logistics cost. J China Railway Soc 31(2):1–6
7. Cao J, Zhu S (1993) Comprehensive compilation of railway online train formation plan. J Southwest Jiaotong Univ 05:93–98
8. Wang L, Ma J, Lin B, Ni S, Lu H (2013) Comprehensive optimization of direct loading and technical direct and section train formation plan. J Transp Syst Eng Inf 13(5):127–133+144
9. Fan Z, Lin B, Niu H (2007) Research on the direct driving plan of the loading place. Logistics Technol 06:77–80

Research on Accident Causing Chains with Bayesian Networks on Waterborne Engineering

Junyong Wang and Yongrui Wen

Abstract The accident mechanism is established to strengthen the safety management and reduce accidents of waterborne transport projects combined with security science. Basing on "2-4" model theory, this mechanism analyzes the influences on the security of waterborne projects construction. The impact factors include personal behaviors, project management and environment. We calculate the probability of each factor with Bayesian network and rank them as they contribute the accident. According to the probability, prevention measure and security management methods can be made to reduce accidents.

Keywords Waterborne engineering · Accident mechanism · Bayesian network

Accident causation theory has developed since 1930s. It is divided into three periods: the classical, early modern and modern. In the classical theory, objects are considered as the major factors causing accidents; in the early modern theory, research on the management factor is made; the modern theory takes human behavior as the major factor causing accidents [1].

In recent years, waterborne transportation has developed rapidly. The management of waterborne projects is especially important to ensure the safety of construction of waterborne infrastructure, as accidents happen often [2]. So far, existing accident causation theories concentrate mainly on the coal industry without taking into account the complex environmental condition and management factors [3]. A well-developed theory has not yet been constructed to reveal the accident factors in the waterborne projects construction [4].

This paper is based on National Key R&D Program: 2017YFC0805300.

J. Wang (✉) · Y. Wen (✉)
China Waterborne Transport Research Institute, Beijing 100088, China
e-mail: wjychqyj@163.com

Y. Wen
e-mail: wenyongrui@wti.ac.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_70

1 Accident Causation Chain

This study shows that accidents are resulted from three kinds of accident causation chains, including human behavior, object and environment chain.

1.1 The Human Behavior Accident Causation Chain

The human behavior accident causation chain causes accidents from two aspects: corporate level and project level. In the corporate level, the corporate safety concept shortness results in corporate management system defect, which further leads to supervision defect [5]. In the project level, the safety concept shortness also results in management system defect, which further leads to human unsafe behavior including command error, operate error and supervise error.

1.2 The Object Accident Causation Chain

The object accident causation chain causes accidents with three major factors: material defect, facility defect and protection defect. These three factors are mainly related to the project safety management system [6].

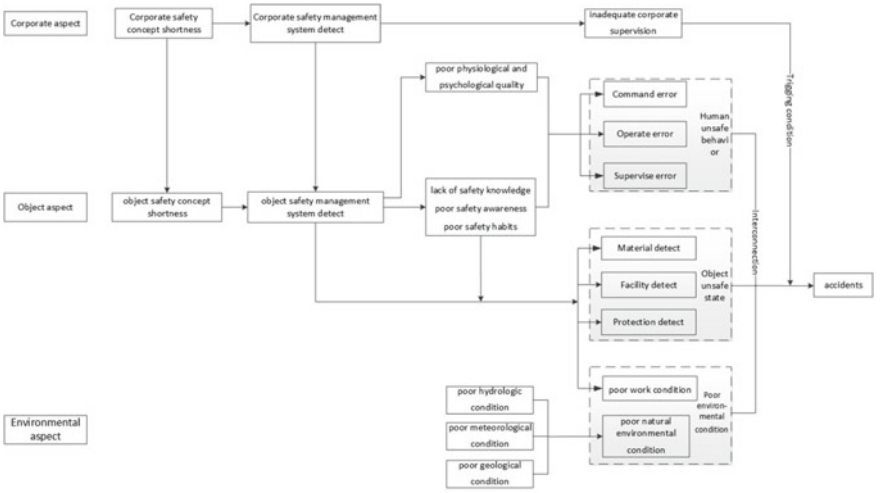
The object unsafe state and human unsafe behavior interact with each other and together lead to accidents.

1.3 The Environment Accident Causation Chain

The environmental factors include natural environmental condition and work condition. The natural environmental condition is related to neither human behavior nor object factors. The work condition depends on the safety management system. Human unsafe behavior, object unsafe state and bad environmental condition together lead to accidents [7, 8].

2 Accident Causing Model

The accident mechanism has been constructed based on the interaction of the accident causation chains [9]. It can show the interaction of human unsafe behavior, object unsafe state and bad environment condition [10]. The model is showed in Picture 1.



Picture 1 Accident causing model

From Picture 1, the accidents of waterborne projects are caused by the interaction of human unsafe behavior, object unsafe state and bad environment condition. The inadequate supervision of waterborne projects is the activating condition of accidents. Good corporate supervision of projects will help discover and change human unsafe behavior, object unsafe state and bad environment condition. Then corresponding measures can be taken to avoid accidents.

Accident-causing factors include corporate, project and environmental factors. In corporate level, the root cause is shortness of corporate safety concept. Leaders concern little about the management of the safety during construction with a fluke mind. The root cause results in the fundamental cause: corporate management system defect, and then lead to the defect of project management system.

In project level, the root cause of accidents is safety concept shortness of project team. The root cause results in the fundamental cause: defect of project management system. Imperfect management system, insufficient training and deficient organization and management lead to safety knowledge shortness, safety consciousness shortness and poor safety habits. These three factors lead to human unsafe behavior: command error, operate error and supervise error.

Study on the analysis of accident-causing mechanism shows that corporate factors result in project factors, while environmental factors are uncorrelated with human behavior or object state. The shortness of corporate safety concepts and management system lead to defect of project team's safety concepts and management system. Leaders' poor supervision of waterborne projects is the activating condition of accidents during construction.

The happening of accidents is the result of interaction of human unsafe behavior, object unsafe state and poor environmental condition. For example, command error

with protection defect will cause accidents. In the same way, operate error with facility defect will also cause accidents.

Human unsafe behavior, object unsafe state and complex environmental condition cause accidents and infect each other. Human unsafe behavior may cause object unsafe state and then cause accidents. Poor environmental condition may directly cause accidents.

3 Bayesian Network

The fault tree model of accident-causing chain is built after the research on the accident cases of waterborne projects construction and the risk estimation of safety of waterborne projects construction.

We do researches on the relationship between all the factors that cause accidents during construction of waterborne projects, and calculate and analyze the contribution of all factors to the accidents.

According to Picture 1, the accident factors can be divided as:

a. corporate safety concept shortness, corporate management system defect and inadequate corporate supervision;
b. object safety concept shortness, object management system defect;
c. poor physiological psychological quality, lack of safety knowledge, poor safety awareness and poor safety habits;
d. human unsafe behavior includes command error, operate error and supervise error;
e. object unsafe state includes material detect, facility detect and protection detect;
f. environmental condition includes poor work condition and poor natural environmental condition which covers poor hydrologic condition, poor meteorological condition and poor geological condition.

All the factors are shown in Table 1.

We turn the accident-causing mechanism into Bayesian network to show the logical relationship among all the factors, the network is as following in Picture 2.

The advantage of Bayesian network is probability updating, when new information observed. The joint probability P(U) in the Bayesian network can be calculated as:

$$\prod_{i=1}^{n} P[X_i|P_a(X_i)] = P(X_3|X_1, X_2)P(X_1)P(X_2)$$

In which:

$U = \{X_1,\ldots,X_n\}$ is a variate;
$P_a(X_i)$ is the super set of the variate X_i;
$P[X_i|P_aP(X_i)]$ is conditional probability of the variate X_i.

Table 1 Factors of waterborne project accident

[Basic elements]					
[Corporate aspect]	[Project aspect]		[Human unsafe behavior]	[Object unsafe state]	[Environmental condition]
[A1 corporate safety concept shortness]	[B1 object safety concept shortness]	[C1 poor physiological and psychological quality]	[D1 command error]	[E1 material defect]	[F1 poor work condition]
[A2 corporate management system defect]	[B2 object management system defect]	[C2 lack of safety knowledge]	[D2 operate error]	[E2 facility defect]	[F2 poor natural environmental condition]
[A3 inadequate corporate supervision]		[C3 poor safety awareness]	[D3 supervise error]	[E3 protection defect]	[F3 poor hydrologic condition]
		[C4 poor safety habits]			[F4 poor meteorological condition]
					[F5 poor geological condition]
					[T accidents]

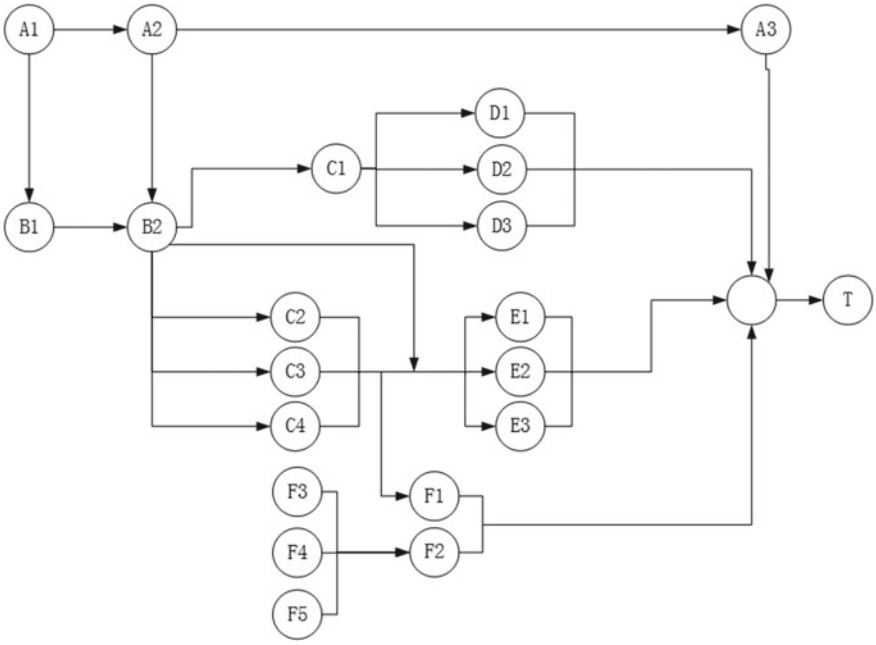
Owing to the characteristic of Bayesian network, the variation of the prior probability causes little error of the probability of final accidents, when the net structure and causal relationship of the nodes are determined. Setting the probability of basic factors as 0.0001, the probability of each factor is calculated and ranked as following:

$$A_1 > A_2 > A_3 > F_3 = F_4 = F_5 > B_1 > B_2 > C_1 > C_2 = C_3 = C_4 > F_2 = F_1 > D_1 = D_2 = D_3 = E_1 = E_2 = E_3$$

The results show the main factors of accidents are A_1, A_2, A_3, F_3, F_4 and F_5.

4 Conclusion

This study analyzes the development of all the accident-causing theories since the 1930s, and does comparative researches on advantages, disadvantages and the range of application of the accident-causing theories. The study shows the blank area of theoretical study that there is no suitable theory for the safety of waterborne projects during construction. Combining safety theories and waterborne projects, this study



Picture 2 Bayesian network of accident factors

constructs an accident-causing chains and mechanism with management factor taken into consideration.

The main causing factors can be selected with the quantitative calculation of the importance degree of all the factors. According to the Bayesian network and calculation, A_1, A_2, A_3, F_3, F_4 and F_5 are the main factors that cause the accidents.

Measures can be taken to control the accidents during the construction of waterborne projects. The fundamental method is to improve the corporate safety concept and to consummate corporate management system. The direct method is to strengthen the corporate supervision, reduce human unsafe behavior, decrease object unsafe state and control the work condition.

The method in this study is used to analyze the main causes of a specific waterborne projects: the caisson structure. The results fit the practical factors of the safety.

This study improves the research method with a combination of existing accident-causing models and waterborne projects. Mathematical method is also used in the study to make quantitative calculation.

The accident-causing mechanism constructed in this study can analyze the main causes in the construction of waterborne projects, and provide prevention measures.

References

1. Fu G, Yin W, Dong J et al (2013) Behavior-based accident causation: the "2-4" model and its safety implications in coal mines. J China Coal Soc 38(7):1123–1129
2. Reason J, Hollnagel E, Paries JS (2006) Revisiting the "Swiss Cheese" model of accidents. Euro Control Experimental Center, Brussels
3. Stroeve SH, Sharpanskykh A, Kirwan B (2011) Agent-based organizational modeling for analysis of safety culture at an air navigation service provider. Reliab Eng Syst Saf 96:515–533
4. Park H (2011) Man-made disasters: a cross-national analysis. Int Bus Rev 20:466–476
5. Hollnagel E (2004) Barriers and accident prevention. Ashgate Publishing Group, Burlington, pp 21–50
6. Qureshi ZH (2008) A review of accident modelling approaches for complex critical socio-technical systems. Defence Science and Technology Organization
7. Reason J (1990) Human error. Cambridge University Press, Cambridge, pp 173–188
8. Fan Y, Guo Y (2014) Csusal factor analysis of Chinese coal mining accident based on HFACS frame. Disaster Adv 7(4):19–26
9. Geoff C, Hussein S (1998) A project specific risk management concept. Int J Proj Manage 16(6):353–366
10. Everson (2004) A systems-theoretic approach to safety engineering. In: Monograph of the ESD Symposium, pp 41–44

Study on the Impact on Drivers of Performance Difference Between Pure Electric and Conventional Fuel Bus

Wei-hua Zhao, Kai-xi Yang, Yu-han Li and Chu-Na Wu

Abstract The accident rate would rise first and then reached the same level when the traditional fuel buses are replaced by pure electric vehicles in batch. In order to study this special phenomenon the impact on bus drivers were studied. The causes were studied to short the adaptation period of drivers and improve the safety level. 200 bus drivers were regarded as the research objects. The actual driving parameters were measured. The bus drivers were investigated. The passengers and traffic participants were investigated too. Driving speed estimation, the acceleration of driving, the driver drowsiness and vehicles found as evaluation criteria were selected as assessment index to compare the impact on the drivers from pure electric bus and conventional fuel bus. Result showed that the driving speed estimation errors were larger from pure electric vehicle in the early stages of the vehicle replacement. The acceleration of driving and braking deceleration were larger from pure electric vehicle. Low labor load from pure electric vehicle caused dull sleepy phenomenon appeared early compared to conventional fuel bus. Pure electric vehicles were more difficult to be found by traffic participants for the low noise. These were the important cause of accidents rate rising early after vehicle replacement. The reasons attributed to driver fatigue and attention distracted were fallacy in operation management. Strengthen the driver management and education could not change this situation. After the replacement of fuel oil buses by pure electric buses, drivers should experience these differences. The driving skills would not be the only requirements.

Keywords Pure electric bus · Performance · The driver · Driving behavior · Drab fatigue

W. Zhao (✉) · K. Yang · Y. Li · C.-N. Wu
Automotive Detection Engineering Technology Research Center, Xi' an Aeronautical University, Xi'an 710077, China
e-mail: xhzwh@outlook.com

Key Laboratory of Operation Safety Technology on Transport Vehicles, Ministry of Transport, PRC, Beijing 100088, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_71

1 Introduction

In the development of electric vehicles in China, urban buses are the pioneers. On July 3, 2018, the State Council issued the "Three-year Action Plan for Winning the Blue Sky Defense War", pointing out that by the end of 2020, all buses will be replaced by new energy vehicles in municipalities directly under the Central Government, provincial capitals and city specifically designated in the state plan. The number of buses in China is now more than 500,000. China's urban buses are operated by large public transport companies, so vehicle renewal is often carried out on a large scale at one time. Compared with the traditional fuel vehicle, the performance of pure electric vehicle has changed greatly, especially in comfort, operation convenience and dynamic response characteristics. Although theoretically, the vehicle performance is continuously optimized, the original vehicle driver's operation is more simple, and the driving safety will be higher. However, after the large-scale replacement of traditional fuel vehicles by pure electric vehicles in many cities, the number of traffic accidents increased within three months, and after about more three months, the phenomenon of high accident rate gradually disappeared. In the process of operation of a bus company, it often comes down to management problems. However, as a common phenomenon after vehicle renewal, there should be an inevitable problem of vehicle adaptation.

2 Literature Review

Walker et al. [1] investigated how task length and fatigue influenced the tendency to mind-wander while driving. They were also interested in whether the propensity to mind-wander could be predicted by individual differences in sustained attention, as measured by the Sustained Attention to Response Task (SART). This research has implications for both basic and applied research on individual differences and cognitive distraction, as well as practical safety implications in areas of driver training and autonomous vehicle development. Nowosielski, RJ et al. have studied distracted driving [2]. Although the research has focused on the deleterious effects of distraction, there may be situations where distraction improves driving performance. Fatigue and boredom are also associated with collision risk and it is possible that secondary tasks can help alleviate the effects of fatigue and boredom. Furthermore, it has been found that individuals with high levels of executive functioning as measured by the OSPAN (Operation Span) task show better driving while multitasking. Yun Meiping et al. quantitatively analyzed the influence of ride comfort and body vibration on ride comfort during bus operation [3]. Using smart phone as data acquisition terminal, the multidimensional variables of ride comfort were constructed by statistical analysis of speed, longitude and latitude data, and the dimensionality of ride comfort was reduced by principal component analysis. Lu et al. studied the time-varying law of driver's alertness in monotonous environment [4]. There is no significant difference

in the driver's heart rate during the driving process, but the degree of subjective fatigue increases significantly. Wang [5] started from the driver's physiological factors,combed the research achievements and deficiencies of driver's behavior and risk perception at home and abroad from three aspects: driver's vision,obstructive sleep apnea syndrome and musculoskeletal disorders. The results show that drivers' physiological illness has a significant impact on their driving ability and risk perception. There is a significant positive correlation between the severity of physiological illness and general and dangerous driving behavior. The higher the severity of illness,the higher the frequency of trip errors. Qin et al. [6] used the driving simulation system platform to test the visual attention needs and driving behavior of 21 subjects in different traffic flow states,using different lane types and different traffic flow states (free/stable/unstable/forced) as virtual test sites,using the psychological test design method. The results show that there is a correlation between driver's visual attention demand and braking times, accelerator pedal displacement and vehicle trajectory deviation. Mutoh et al. [7] studied a new type of electric vehicle (EV) system with independently driven front rear wheels. It has many advantages which are needed for city cars such as good steering ability in congested traffic and complete failsafe functions to guarantee the safety at the time of failure. Numasato et al. [8] presented a settling control of a dual-actuator system for hard disk drives. The dual-actuator system consists of a voice coil motor (VCM) as a first stage actuator and a push-pull-type piezo-electric transducer (PZT) as a second-stage actuator. Experimental results show that the dual actuator system with the proposed settling controller achieves better performance than a single actuator system with the same VCM and a conventional settling controller. Mauro et al. [9] shortly described an ABS/ESP Hardware-In-the-Loop (HIL) test bench built by the Vehicle Dynamics Team of the Department of Mechanics of Politecnico di Torino. It consists of a whole brake system, integrated through specific interface (e.g. wheel pressures signals) with a vehicle model running in real time on a dSPACE board. A brief overview of HIL application for developing an Electro-Hydraulic Braking system (EHB) was provided. Numasato et al. [10] presented manuscript investigates the current sustainability research within the automotive industry, through a comprehensive review of the different studies in vehicles' life cycle, disposal and end of life analyses, and the different sustainability metrics and models used to quantify the environmental impact. Kepner [11] studied Hydraulic Hybrid Vehicle (HHV) technology. The inherent power density of HHV made significant benefits from regenerative braking possible in higher- mass vehicles. Other advances in hydraulic components make HHV practical in a passenger vehicle. Oshima et al. [12] presented an electrically driven intelligent brake system that has been developed for electric vehicles and hybrid electric vehicles and outlined the newly developed brake system and describes various issues involved in cooperative regenerative braking along with the technologies that were applied to address them. George et al. [13] studied twelve male subjects who rated the discomfort caused by lateral oscillation at eight frequencies (0.2–1.0 Hz) across four seating conditions (a rigid seat and a train seat, both with and without backrests) and showed that low frequency lateral acceleration can cause less discomfort when sitting with a backrest than when sitting on the same seat without a backrest. Pan

et al. [14] proposed a braking intention classification method to improving comfort. And result showed that the braking comfort and safety can be improved by using the proposed braking intention classification and recognition method. Lu [15] using pure electric vehicle as the research object, from the perspective of ergonomics, the comfort analysis, evaluation and improvement of pure electric vehicle are studied and discussed. However, there is a lack of research on drivers of pure electric buses in the literature. For drivers, driving fuel buses and pure electric buses must be different, so this paper would study and analyze the differences from four aspects.

However, there is a lack of research on drivers of pure electric buses in the literature. For drivers, driving fuel buses and pure electric buses must be different, so this paper will study and analyze the differences from four aspects.

3 Research Methods and Experiments

3.1 Research Methods and Objects

This paper chooses a bus company in a city of China as the research object, investigates 200 bus drivers who invested in the previous period, summarizes the main safety problems before and after vehicle renewal, and analyses the possible reasons. Based on the main reasons of driver survey, parameters are set. Data monitoring and analysis were carried out on 100 newly invested pure electric buses and the related performance of replacing traditional buses. The differences between pure electric buses and traditional fuel buses were compared, and the reasons for the problem analysis were determined.

3.2 Test Equipment

(1) Vehicle speed and acceleration and deceleration parameters are measured by speedbox to determine the changing law of vehicle driving state.
(2) Using the noise meter to measure the noise inside and outside the vehicle, to determine the noise inside and outside the driver's driving environment.
(3) Using stopwatch to measure the driver's drowsiness occurrence time.
(4) By using the speed box speed monitoring function, the driver is asked by the subject at the set speed, and the driver estimates the driving speed, which is deviated.

3.3 Test Method

(1) Vehicles are driven according to the routine driving habits on the running line, and the driving speed is observed and recorded by the main test. The set observation speeds are 0, 10, 20, 30, 40 and 50 km/h respectively. The driver is asked to record the estimated speed without observing the speedometer.
(2) In the course of driving, the acceleration of the vehicle is recorded by speedbox, including the start acceleration and the brake acceleration.
(3) During the driving process, the driver, according to his own state changes, notifies the main test to record the occurrence time of sleepiness when the feeling of sleepiness occurs.
(4) Investigate the participants in vehicle traffic, and investigate the situation of electric motorcycles and drivers finding buses while driving, to study the possibility of electric buses being discovered by other vehicles.

4 Experimental Result

4.1 Speed Estimation

4.1.1 Difference Test

According to different driving speeds, the speed estimation results of all subjects driving different vehicles were classified and counted, and the speed estimation data were obtained. All the data were input into SPSS for statistical analysis. A paired T test is performed for the speed estimation results of the same driver at different speeds. The differences between the speed estimation results of traditional fuel vehicles and electric vehicles are tested. The results are shown in Table 1.

The test results show that there is a significant difference in speed estimation between traditional fuel vehicles and pure electric vehicles.

Table 1 Significance of speed estimation differences for different vehicle types

Speed (km/h)	10	20	30	40	50
Test result	0.043	0.007	0.006	0.015	0.029

*There was a significant difference in speed estimation between traditional fuel vehicles and pure electric vehicles ($P < 0.05$)

4.1.2 Change Regulation of Speed Difference Estimation

During the driving process, the driver will estimate the speed according to the scene switching speed and the change of engine sound. When driving a vehicle for a long time, the driving speed is more dependent on the change of engine sound. However, with the use of pure electric vehicles, the low-level noise environment causes the loss of driver speed estimation information, so the accuracy of speed estimation decreases. Because only the accuracy of speed estimation is considered, the difference between the estimated value and the real driving speed is taken as absolute value, which can measure the accuracy of speed estimation. The specific characterization is shown in Formula 1.

$$D = |S_e - S_r|$$

Among them, D is the difference of speed, km/h; S_e is the estimation of speed, km/h; S_r is the real speed, km/h.

Under different actual driving speeds, the statistical data of the difference between pure electric bus and traditional fuel vehicle speed estimates are shown in Table 2. The difference distribution of pure electric bus speed estimation is shown in Fig. 1, the difference distribution of traditional fuel vehicle speed estimation is shown in Fig. 2, and the difference between traditional fuel vehicle and pure electric vehicle varies with driving speed as shown in Fig. 3.

From Table 2, it can be seen that the greater the speed, the greater the difference in the average estimated speed, that is, the greater the speed, the more inaccurate the speed estimation. For the same speed, the average error of pure electric vehicle speed estimation is greater than that of fuel vehicle.

After data processing, it can be seen that the speed estimation error of traditional fuel vehicles is less than that of pure electric vehicles. The variation law of speed estimation difference of traditional fuel vehicles is shown in Formula 2, and that of pure electric buses is shown in Formula 3.

$$D_f = y = 3.3779 \ln(x) + 3.8577$$

Among them, D_f is the difference of estimated speed of traditional fuel vehicles, km/h; S_r is the real speed, km/h.

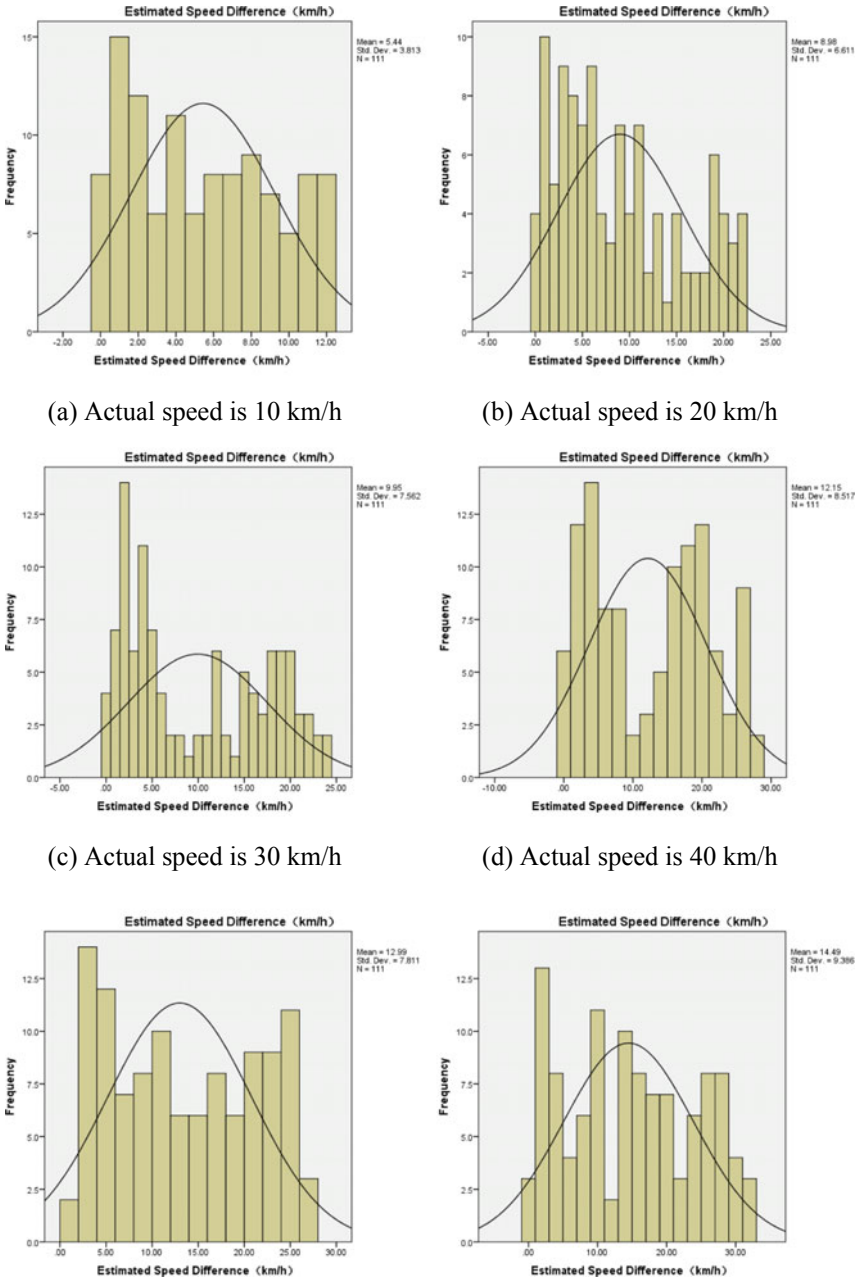
$$D_e = y = -0.1914x^2 + 3.0394x + 2.932$$

Among them, D_e is the estimated speed difference of pure electric vehicle, km/h; S_r is the real speed, km/h.

As can be seen from Fig. 3, the median of the difference between the two kinds of vehicle speed estimates is basically less than the average, and increases with the increase of vehicle speed. But the error of speed estimation of pure electric vehicles is larger than that of traditional fuel buses.

Table 2 Statistical data of the difference between pure electric bus and traditional fuel vehicle speed estimates

Actual speed	10 km/h		20 km/h		30 km/h		40 km/h		50 km/h		60 km/h	
Type	Pure electric bus	Traditional fuel bus	Pure electric bus	Traditional fuel bus	Pure electric bus	Traditional fuel bus	Pure electric bus	Traditional fuel bus	Pure electric bus	Traditional fuel bus	Pure electric bus	Traditional fuel bus
Average value	5.44	3.32	8.98	7.42	9.95	7.58	12.15	8.00	12.99	8.38	14.49	10.67
Average standard error	3.81	2.23	6.61	4.39	7.56	5.72	8.52	5.37	7.81	5.34	9.39	7.90

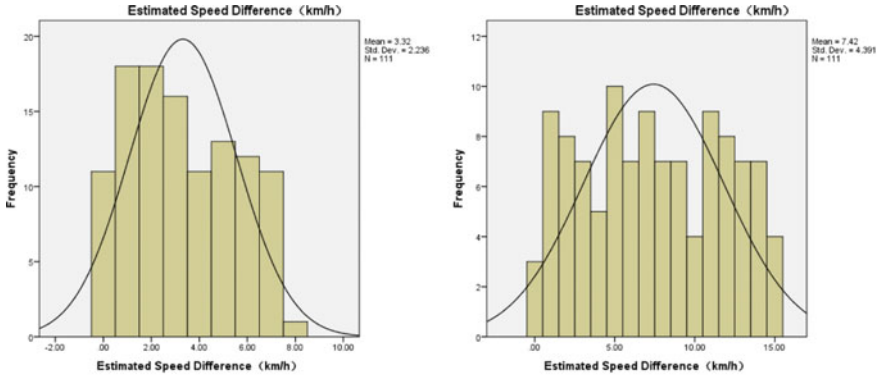


(a) Actual speed is 10 km/h (b) Actual speed is 20 km/h

(c) Actual speed is 30 km/h (d) Actual speed is 40 km/h

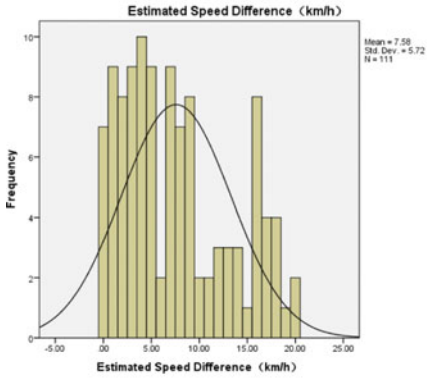
(e) Actual speed is 50 km/h (f) Actual speed is 60 km/h

Fig. 1 Difference distribution of velocity estimation for pure electric bus at different vehicle speed

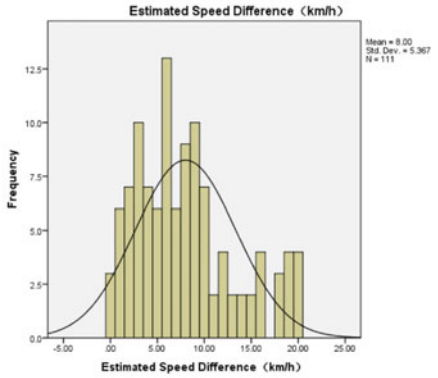


(a) Actual speed is 10 km/h

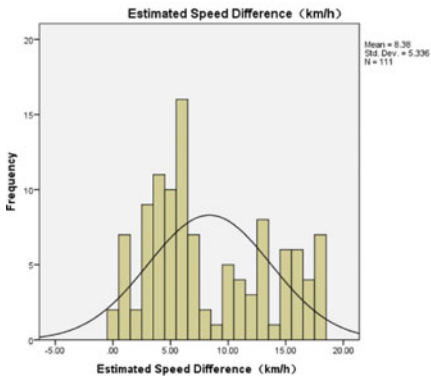
(b) Actual speed is 20 km/h



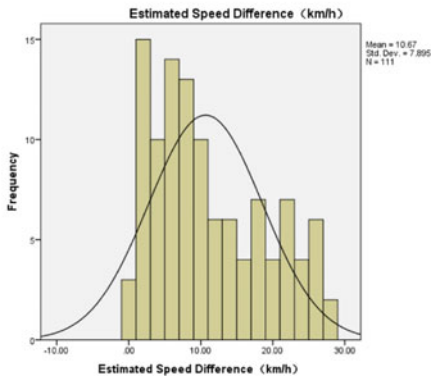
(c) Actual speed is 30 km/h



(d) Actual speed is 40 km/h



(e) Actual speed is 50 km/h



(f) Actual speed is 60 km/h

Fig. 2 Difference distribution of velocity estimation for traditional fuel bus at different vehicle speed

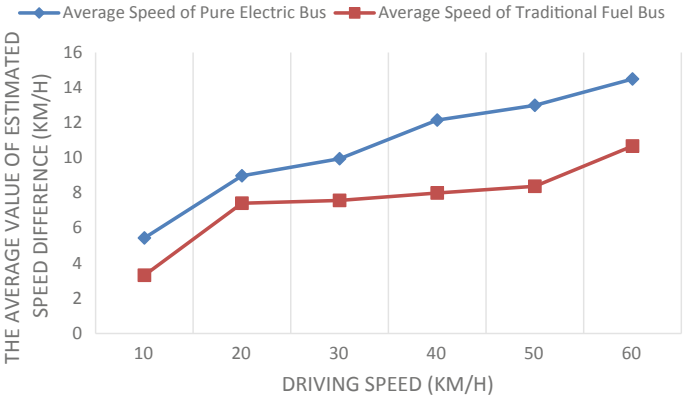


Fig. 3 The average value of estimated speed difference between conventional fuel vehicles and pure electric vehicles varies with driving speed

4.2 Driving Acceleration

Traffic acceleration is mainly aimed at starting acceleration and braking acceleration to analyze the two indicators, to characterize the comfort difference between the two vehicles. Because the acceleration of the bus is affected by the distance between the road traffic scene and the obstacle in front of it, the acceleration obtained by using the speedbox monitoring can not accurately describe the change of the driving state of the vehicle. Therefore, we use the method of driver and passenger survey to get relevant conclusions. The subjects included 100 bus drivers and 236 passengers on different vehicles. The driver's survey results are shown in Table 3 and passengers' feelings are shown in Table 4.

From the results of the investigation in Tables 3 and 4, it can be seen that the common characteristics of buses are frequent speed change, high starting and braking acceleration. From the driver's point of view, after getting used to the performance

Table 3 Driver survey results

Survey items	Pure electric bus		Traditional fuel bus	
	Subjective feeling	Accounting situation (%)	Subjective feeling	Accounting situation (%)
Acceleration frequency	High	73.25	High	23.45
Acceleration	High	87.13	High	22.51
Frequent braking	High	86.26	High	53.21
Braking acceleration	High	65.64	High	10.43

Table 4 Survey results of ride comfort

Survey items	Pure electric bus		Traditional fuel bus	
	Subjective feeling	Accounting situation (%)	Subjective feeling	Accounting situation (%)
Acceleration frequency	High	98.25	High	85.46
Acceleration	High	97.85	High	86.54
Frequent braking	High	96.29	High	76.62
Braking acceleration	High	96.66	High	72.68

of the traditional fuel vehicles, people with different identities feel little about the speed change process, but they feel obvious about the speed change process of the electric bus. From the passenger's point of view, the frequent speed change and high acceleration in the process of riding make the ride comfort of electric vehicles worse. Although the acceleration and deceleration characteristics are related to traffic characteristics, the fast acceleration and deceleration response of pure electric vehicles is the direct cause of poor ride comfort.

4.3 Driver Sleepness Occurrence Time

Compared with traditional fuel vehicles, pure electric bus lacks frequent operation of clutch and shift rod, and power system lacks vibration and noise of engine. Corresponding to this change, the complexity of driving operation is greatly reduced, and the noise of interior power system is greatly reduced. According to the measurement results of vehicle interior noise, the interior noise of pure electric bus in static state is 33–52 min DB, while the idle noise of traditional fuel vehicle is 58–67 DB. When a driver suddenly changes from a traditional fuel bus to a pure electric bus, the driver's subjective feeling is that the operation is too simple and too quiet in the same traffic scenario, and the long-term attention allocation habits of the driver will change. The comparison results of driving operation differences are shown in Table 5.

From Table 5, it can be seen that bus drivers are more skilled than private drivers because they are familiar with the operation of a fixed line for a long time. Even if they drive a fuel car, they do not find it difficult to operate. But in the initial stage of converting to electric vehicle, because no relevant operation is needed, the operation complexity is no operation for professional drivers, which makes the driver pay little attention to the operation. At the same time, the interior noise is low, so it is difficult to feel the vibration of the power system, which makes the driver in a relatively relaxed state. In monotonous environment and operation, drivers are prone to sleepiness. Statistical analysis of the occurrence time of drowsiness shows that the driver's response is shown in Table 6 when driving traditional fuel vehicles and pure electric vehicles.

Table 5 Results of Driving Operations Difference Survey

Survey items	Pure electric bus		Traditional fuel bus	
	Subjective feeling	Accounting situation (%)	Subjective feeling	Accounting situation (%)
Shift operation complexity	Uncomplicated	100	Generally complex	100
Clutch operation complexity	Uncomplicated	100	Generally complex	100
The environmental noise inside the car feel	Low	100	High	100

Table 6 Statistics of sleeping time of drivers of pure electric buses and fuel buses

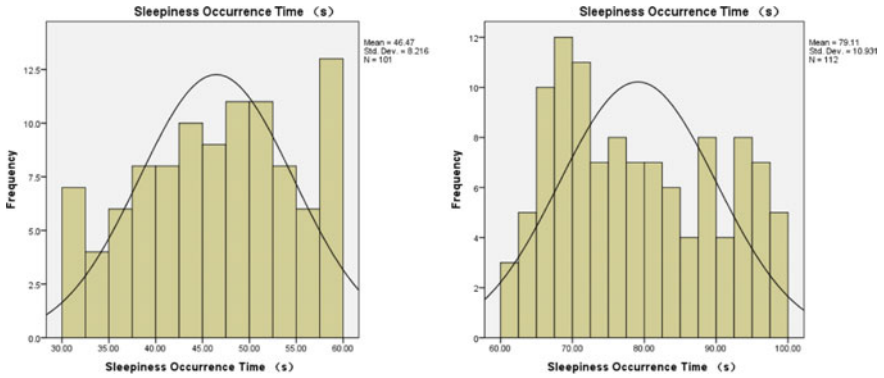
	Pure electric bus	Traditional fuel bus
Average value	46.47	79.11
Average standard error	0.818	1.033
Median	47.38	77.43
Mode number	32a	67
Standard deviation	8.216	10.931
Variance	67.502	119.489
Minimum value	31	61
Maximum value	60	100

Table 6 shows that the average time and median time of sleepiness of pure electric bus drivers are much shorter than that of fuel buses, and the overall range of sleepiness time is also smaller than that of fuel buses.

As shown in Fig. 4, the requirements for drivers of pure electric buses are greatly reduced due to the monotonous driving operation, comfortable driving environment and good silence effect. In addition, bus drivers are particularly familiar with the driving routes and have low attention tension. The phenomenon of sleepiness occurs much earlier than that of traditional fuel vehicles. This sleepiness is not caused by fatigue, but by the monotony of the environment.

4.4 Sensory Distance of Traffic Participants

In actual traffic scenarios, various sound information generated by other vehicles is the main way for traffic participants to perceive vehicle information on the road. Because pure electric bus has no engine noise, the vehicle technical condition is better, which leads to a significant reduction of noise in the course of driving. Compared



(a) Pure Electric Bus (b) Traditional Fuel Bus

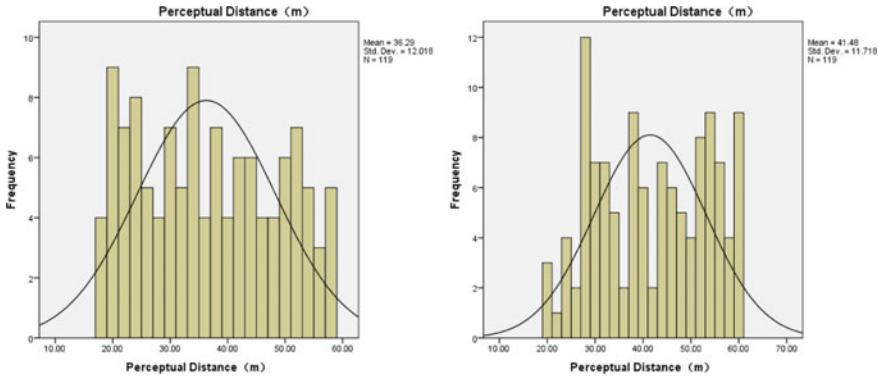
Fig. 4 Distribution chart of sleepness occurrence time for drivers of pure electric bus and fuel bus

with the actual measurement, the noise of traditional fuel vehicles is 78–103 dB, while that of pure electric buses is 56–87 dB. There is little difference between the noise in this area and the background noise in the urban environment, which makes it difficult for the traffic participants to obtain the driving information of the vehicle and easy to cause traffic accidents. By investigating different types of traffic participants on the road, it is found that the perceived distance of pure electric bus approaching information is significantly different from that of traditional fuel vehicles. The specific results are shown in Table 7.

From Table 7, it can be seen that the pedestrian discovery distance of pure electric bus is 5.2 m shorter than that of traditional fuel bus, 8.5 m shorter than that of traditional fuel bus, 13.3 m shorter than that of motorcycle driver, and the pedestrian discovery is earlier than that of motorcycle driver as a whole, motorcycle discovery is earlier than that of motorcycle driver, while motorcycle driver discovery is the latest, which conforms to common sense (Figs. 5, 6 and 7).

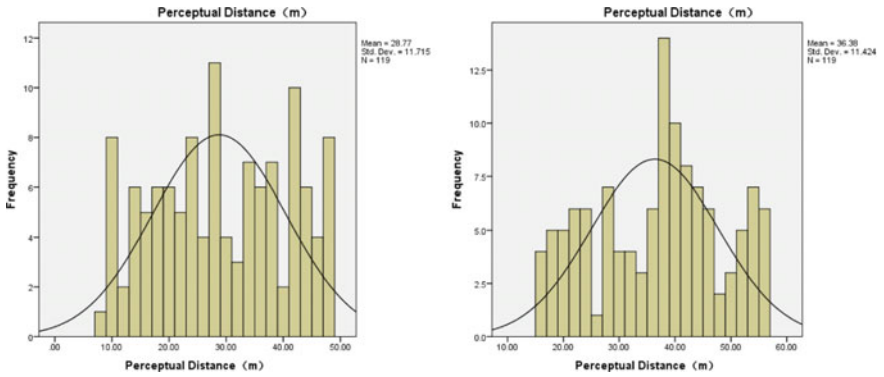
Table 7 Distance statistics of pedestrians, motorcycle drivers and automobile drivers finding pure electric buses and fuel buses

Research object	Pedestrian		Motorcycle		Automobile	
Test object	Pure electric bus	Traditional fuel bus	Pure electric bus	Traditional fuel bus	Pure electric bus	Traditional fuel bus
Average value	36.2857	41.4790	28.7731	36.3782	16.1345	29.8571
Average standard error	1.10166	1.07415	1.07393	1.04724	0.95415	1.06044



(a) Pure Electric Bus (b) Traditional Fuel Bus

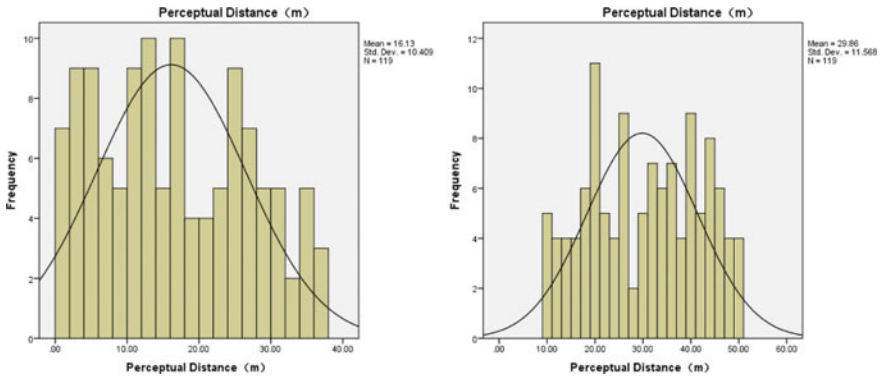
Fig. 5 Distance distribution of pedestrian discovery vehicle



(a) Pure Electric Bus (b) Traditional Fuel Bus

Fig. 6 Distance distribution of motorcycle drivers discovering vehicles

Due to the narrowing of the detection distance, the reaction time left to traffic participants is short after the pure electric bus approached, which is easy to induce accidents. Through the investigation of bus drivers, it is also verified that traffic participants have low sensitivity to pure electric bus approaching information, and the risk level of traditional fuel bus driving habits is high, which leads to sudden increase of braking during driving.



(a) Pure Electric Bus (b) Traditional Fuel Bus

Fig. 7 Distance distribution of vehicles discovered by drivers

5 Conclusion

Through the analysis of the data, it is found that the main reasons for the increase of accidents are the driver's drowsiness, inaccurate estimation of driving speed, excessive starting acceleration and braking deceleration, and the traffic participants' inability to be alert when the traditional fuel vehicles are replaced by pure electric buses. After the occurrence of this phenomenon, bus operators often attribute it to driver fatigue, inattention and other reasons, and then strengthen the management and education of drivers, ignoring the root causes of the problem. The actual operation statistics of enterprises show that the accident rate of drivers will decrease dramatically after about three months, which also proves this problem. Therefore, after the pure electric bus replaces the fuel bus, the driver needs to go through a period of adaptation period, and in this period, the work undertaken should be adjusted accordingly to avoid accidents.

Acknowledgements Supported by the Opening project of Key Laboratory of Operation Safety Technology on Transport Vehicles, Ministry of Transport (KFKT2018-05).

References

1. Walker HEK, Trick LM (2018) Mind-wandering while driving: the impact of fatigue, task length, and sustained attention abilities. Transp Res Part F Psychol Behav 59:81–97
2. Nowosielski RJ, Trick LM, Toxopeus R (2018) Good distractions: testing the effects of listening to an audiobook on driving performance in simple and complex road environments. Accid Anal Prev 111:202–209
3. Yun M, Wang W (2017) Smartphone based research of measurement indexes related to bus riding comfort. J Tongji Univ (Natural Science) 45(8):1143–1149

4. Lu Q, Sun Y, Zhu T (2017) The variation of drivers' vigilance with time under the monotonous environment. Technol Innov Manage 38(01):36–40+109
5. Wang C (2018) Review of research on effect of physiological factors on behavior and risk perception of drivers. J Saf 14(4):155–159
6. Qin Y, Hui Y, Guo F, Li H, Zhang Jing (2017) Analysis of the relation between the visual attention demand and the motivation characteristics of the driving behavior. J Saf Environ 17(4):1376–1381
7. Mutoh N, Takita K (2004) A control method to suitably distribute electric braking force between front and rear wheels in electric vehicle systems with independently driven front and rear wheels. In: The 39th IAS annual meeting
8. Numasato H, Tomizuka M (2003) Settling control and performance of a dual-actuator system for hard disk drives. IEEE ASME Trans Mech 4:431–438
9. Velardocchia M et al (2005) Hardware in the loop to evaluate active braking systems performance. SAE J
10. Numasato H, Masayoshi T (2003) Settling control and performance of a dual-actuator system for hard disk drives. IEEE ASME Trans Mech 8(4):431–438
11. Kepner RP (2002) Hydraulic power assist-a demonstration of hydraulic hybrid vehicle regenerative braking in a road vehicle application. SAE Paper, 2002-01-3128
12. Oshima T, Fujiki N, Nakao S (2011) Development of an electrically driven intelligent brake system. SAE J 4:399–405
13. Beard GF, Griffin MJ (2012) Discomfort during lateral acceleration: influence of seat cushion and backrest. Appl Ergon 44:588–594
14. Pan N, Yu L, Song J (2016) Braking intention classification and identification considering braking comfort for electric vehicles. J Tsinghua Univ (Science and Technology) 56(10):1097–1103
15. Lu Z (2015) The study of electric vehicle comfort based on ergonomics. Southeast University

Application of Accident Causation Chain in Security Management of Ports and Channels

Majing Lan and Zhiqiang Hou

Abstract With the rapid development of globalization, China has experienced constantly accelerated development of market economy after joining the WTO, leading to more frequent foreign trade. The port and channel, as important infrastructure, have also expanded its development space. In this paper, the author mainly discusses the application of accident causation chain in security management of port and channel construction by studying the causation chain of classical, modern and contemporary accidents in combination with the characteristics of port and channel construction.

Keywords Accident causation chain · Port and channel · Security management · Application

1 Introduction

With the rapid development of China's maritime trade, China's maritime shipping industry has experienced rapid development. The port is a transportation hub for water and land transportation and channel is an important guarantee for normal operation of water transportation, which play an irreplaceable role in economic development [1]. Therefore, it is imperative to carry out security production management of port and channel construction process in order to protect the security of laborers in the labor process and protect the interests of construction units involved in port and channel projects.

M. Lan (✉) · Z. Hou
China Waterborne Transport Research Institute, Beijing 100088, China
e-mail: Lanbaoshan112@163.com

Z. Hou
e-mail: zhiqianghou@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_72

2 Accident Causation Chain

Generally, the accident causation chain has undergone three stages: classical accident causation chain, modern accident causation chain and contemporary accident causation chain [2]. Although researches on accident causation chain of three stages are conducted in different eras of different productivity levels, they contain different opinions on the direct, indirect and basic causes of accidents, which therefore are of deep practical guiding significance for improvements in the security management level of ports and shipping channels construction [3].

1. **Classical Accident Causation Chain**

In the scope of classical accident causation chain, typical accident causation chain theories include Greenwood and Woods's accident prone tendency theory, Minz and Bloom's Accident Liability, Heinrich's accident causation chine theory and Gordon's accident causation chine theory.

(1) Greenwood and Woods's accident prone tendency theory. Greenwood and Woods believe since a small number of workers are prone to accidents, their existence is the main cause of industrial accidents. Therefore, the incidence of industrial accidents may be reduced by cutting down the number of workers prone to unexpected accidents. In accident prone tendency theory, human personality is blamed for accidents on, which indeed encourages future generations to continue the researches on this subject although it is too simple.
(2) Minz and Bloom's Accident Liability. Mintz and Bloom believe that accidents may occur when various factors such as working conditions, personal characteristics, experience and skills are achieved according to specific trajectories. With the emergence of this theory, people gradually improve focus of security production from strengthening worker management to improving production conditions.
(3) Heinrich's Accident Causation Chain. It can be found from Heinrich accident causation chain theory that accidents always result into injuries, whose direct cause lies in the unsecure behaviors of human and unsecure state of objects, indirect cause lies in shortcoming of human and basic cause lies in the social environment for growth and genetic factors. There is positive significance of this theory. On the one hand, a series of causes and consequences of accidents are connected to form a complete accident causation chain, which offers a line for accident prevention. On the other hand, it offers a prevention method for some accident, which is to eliminate unsecure behaviors and unsecure states, direct causes of accidents. Despite all shortcomings, like the inaccurate indirect and basic causes for accidents, Heinrich's accident causation chain can still be an important theoretical basis for security management.
(4) Gordon's Accident Causation Chain. Gordon believes that accidents are caused by personal characteristics of accidents as well as multifaceted factors influencing individual characteristics. Gordon proposes that the environmental factors of human, media factors close to human and personal characteristics of human

are causes of accidents, which greatly expand the research content of accident causation chain.

2. Modern Accident Causation Chain

In modern accident causation chain, Wiggsworth's accident causation chain as well as Bode and Loftus accident causation chain is typically used for analysis and research.

(1) Wiggsworth's accident causation chain. According to Wiggsworth's accident causation chain, human tend to easily make faults and mistakes in work as a result of their lack of knowledge and education and these faults and mistakes may lead to accidents. According to this theory, accidents are caused by the lack of knowledge and education in management of human. The research that the individual characteristics of human cause various accidents has come as great progress compared with researches prior to the emergence of modern accident causation chain.
(2) Bod and Loftus's accident causation chain. After years of research on accident causation chain, Bode and Loftus first include management factors to factors causing various accidents and shift their research focus from personal characteristics to management compared with previous researches on accident causation chain and make comprehensive consideration of causes of accidents. They analyze the basic causes of unsecure behaviors and unsecure states from the perspective of management factors and reduce the occurrence of accidents by strengthening the improvement of problems in management.

3. Contemporary Accident Cause Chain

So far, contemporary accident causation chain has been dominated by Stewart's modern accident causation chain and behavioral security "2-4" model.

(1) Stewart's modern accident causation chain. Compared with previous researches, Stewart's researches on accident causation chain come as huge progress. On the one hand, Stewart proposes that researches on accident causation chain should be carried out from various aspects such as responsibility of various departments on security, hardware facilities, work quality of security personnel, participation of employees and training status. On the other hand, he also proposes that accident causation chain theory should be studied based on the management of thoughts and activities. In Stewart's accident causation chain theory, the basic cause is management's thinking and activities and the indirect cause is the middle departments and equipment. The source of security performance is the basic cause, while the driving force of security performance is the indirect cause.
(2) Behavioral security "2-4" model. The behavioral security "2-4" model is a modern accident causation chain proposed on the basis of the accident causation chains by Heinrich and by Stewart, where the direct cause is still the wrong one-time behavior of the person causing accidents as proposed by Heinrich while the indirect cause is the habitual behaviors from wrong security habits, security knowledge and security awareness of the person causing accidents. In

this model, the basic cause of accidents is embodied in the lack of security management system of the organization where the accident is triggered; the root cause of accidents is embodied in the lack of security culture of the organization where the accident is triggered. When the "2-4" model is used to analyze accidents, it is possible to get the internal and external factors of accidents in addition to behavioral causation chain of accident initiator. Targeted solutions and prevention suggestions can be proposed according to causes of accidents to effectively prevent the recurrence of similar accidents.

After years of development and research, there has been continuous progress in the research of accident causation chain. Firstly, the analysis and description of causes of accidents are completed from personal characteristics of initiators or direct physical causes of accidents. In addition, the management factor is introduced into accident causation chain as the basis cause. Finally, management factors are divided into several types, which provide a better way for practical accident prevention as well as clearer and more specific practical operation methods.

3 Security Management for Construction in Ports and Channels

Through studying classical, modern and contemporary accident causation chains, it can be found that main causes of accidents are individual behaviors, organizational behaviors and external factors. Therefore, we should focus on management from these three aspects and take measures in advance to eliminate causes of accidents, thereby preventing accidents during the safety management of construction of ports and channels.

1. **Management of Personal Behaviors**

Management of personal behaviors needs to be carried out in the following aspects:

(1) Security inspection. It is necessary to correct habitual behaviors and organizational behaviors in the construction of ports and channels through the supervision department's law enforcement activities and internal security inspection activities, discover unsecure operations of construction workers through discovery, supervision, and correction activities and correct these operation immediately.

(2) Security Education. It is necessary to strengthen the training and education on professional skills and security awareness of construction workers for ports and channels so that construction personnel will standardize operations and operate equipment and facilities with caution in terms of professional skills as well as will realize the importance of security production through security knowledge promotion and warning signs In terms of security awareness. In this way, construction workers will learn about possible security problems and prevention methods.

2. **Management of Organizational Behaviors**

(1) It is necessary to learn about laws and regulations suitable for the company, regularly establish corresponding rules and regulations and management standards in accordance with changes in relevant laws, regulations and norms and make timely promotion and education on relevant requirements so that staff members can have access to relevant knowledge as soon as possible.
(2) It is necessary to formulate and timely revise security operation procedures for each position and issue them to each post.
(3) It is necessary to set up systematic security management system to reduce possible risks, such as clear security management policy, security management organization structure with clear division of function and security plans. All employees should strictly follow the policy of double duties for single post and be responsible for security production within their scope of business.
(4) It is necessary to establish comprehensive emergency plans and special plans to prevent and deal with accidents such as emergency plan for prevention of fire and explosion accidents, emergency plans for prevention of strong wind and typhoon, emergency rescue plans for offshore construction, emergency rescue plans for construction ship overturn, emergency rescue plans for drowning accident, emergency plan for prevention of electric shock accidents and so on [4].

3. **Management of External Factors**

The security management of port and channel construction is under great influence of factors like construction conditions, hydrology, weather and construction vessels on shipping channels. Therefore, construction units should fully consider the natural conditions of construction area, such as typhoons, waves, and currents. The construction units shall regularly organize or entrust relevant agencies for surveys on the meteorological and hydrological conditions in relevant sea area and collect accurate marine climate and hydrological information [5]. In the construction process, the construction unit should attach great importance to security to avoid construction work under harsh environmental conditions, ensure the security of vessels, equipment and personnel during the construction period and take relevant management measures to avoid accidents.

4 Conclusion

Driven by the reform and opening up policy and economic globalization, the ports in China, as an important node in the connection channel with international communities, have been the basis for joining in economic globalization and social division of labor. Therefore, construction of ports and shipping channels has attracted a lot of attention as the backbone of port economic development. Due to the strong professionalism and severe construction conditions for port and channel construction,

there are certain problems in security management which bring potential security hazards in production and work. Regardless of the different causes of various safety hazards and accidents, they share a certain relationship with imperfect management. It is possible to specifically analyze cause of accidents and improve security management with the accident causation chain theory so as to ensure the security of port and channel construction.

References

1. Yang J (2005) Research on theory and practice of modern port development. Shanghai Maritime University
2. Mi F, Yu J, Li S (2014) Analysis and comparison of accident causation chain. China Public Saf (Academic Edition) (01):41–44
3. Su Q (2014) Application of accident causation theory to construction security management. Water Resour Hydropower Constr 1:102–103
4. Zhu X (2016) Research on adaptability of security guarantee of water transportation in Dongying port. Dalian Maritime University
5. Liu Z (2016) Research on risk assessment and emergency countermeasures for oil spills in Long Island Waters. Dalian Maritime University

Comparative Study on the Measures to the Safety Management of Bulk Liquid Dangerous Goods Storage in Port Areas

Chaoyu Ruan, Xin Lu and Zhiqiang Hou

Abstract In this work, the international and Chinese standards of bulk liquid dangerous goods terminal storage and handling are compared and evaluated. The safety management of ship operation in port and the safety management of terminal operation are selected as the key contrasting indexes of the port operation of bulk liquid dangerous goods. Afterwards, a revised proposal for the safety management of bulk liquid dangerous goods in port is proposed.

Keywords Bulk liquid · Dangerous goods · Terminal storage · Safety management

1 Introduction

The different packaging forms of dangerous goods determine the difference in loading and unloading processes used in the port. According to the different types of dangerous goods, the port dangerous goods operations can be divided into: bulk dangerous goods port loading and unloading operations and packaging dangerous goods port loading and unloading operations. In the formulation of port operation practices and standard specifications at home and abroad, the port operations of bulk dangerous goods are usually focused on port operations of bulk liquid dangerous goods. The port loading and unloading of dangerous goods is usually focused on port operations of dangerous goods containers. In this paper, in the handling of dangerous goods handling, the port operations of bulk liquid dangerous goods were selected as key comparative research indicators.

C. Ruan (✉) · X. Lu (✉) · Z. Hou (✉)
Safety and Emergency Technology Center, China Waterborne Transport Research Institute, Beijing, China
e-mail: ruanchaoyu@wti.ac.cn

X. Lu
e-mail: luxin@wti.ac.cn

Z. Hou
e-mail: zhiqianghou@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_73

Internationally, the International Maritime Organization (IMO) proposed the "Proposal for the Transport of Dangerous Goods and Related Activities in the IMO Port Area" (hereinafter referred to as the "Recommendation") in 2007. On the basis of following the relevant international conventions such as the IMDG code, Suggested safety requirements and requirements for dangerous goods in ports operating in bulk liquid dangerous goods. At the same time, the proposal proposes that the port loading and unloading operation of bulk liquid dangerous goods should also meet the requirements of "International Safety Guide for Oil Tankers and Terminals: ISGOTT" and "Liquefied Gas Handling Principles on Ships and in Terminals: LGHP".

China's dangerous goods port operations safety management is based on the "Port Dangerous Goods Safety Management Regulations" as the main line, supplemented by the "Standard Requirements for Safety Technology of Oil Terminals", "Technical Requirements for Port Handling of Bulk Liquid Chemical Products", "General Rules for Port Storage of Bulk Petroleum and Liquid Chemical Products", and "Safety Operation Rules for Oil Tanker Terminals". They together constitute the safety management standard system for dangerous goods port operations in China.

Internationally, considering the continuity and coordination of bulk liquid dangerous cargo ships and shore loading and unloading, ship handling and port loading and unloading and storage are usually considered as two parts of a whole. Therefore, this paper selects the safety management of ships in port operations and the safety management of terminal operations as key comparison indicators for port operations of bulk liquid dangerous goods.

2 Comparison of Safety Management Indicators for Ships Operating in Port

China's safety management requirements for loading and unloading operations of bulk liquid dangerous goods vessels in Port are basically the same as those in the international arena, "Safety Operation Rules for Oil Tanker Terminals" (GB18434-2001) [1] and "International Tanker and Oil Terminal Safety Guide (5th Edition)" (ISGOTT) [2] has put forward requirements for the safe operation of oil tankers in the areas of arrival, berthing, loading and unloading, ballasting, washing, degassing and entering closed compartments, mainly including the following aspects.

A. *Ship internal safety management requirements*

Dangerous goods based on bulk liquids are flammable, explosive, easy to evaporate, easy to generate and accumulate static charges, easy to flow and spread. In China and the world, fire prevention, anti-static and anti-pollution are proposed safety measures in the following aspects:

Smoking and open fire requirements;
Kitchen safety management;

Safety management of electrical equipment;
Thermal work safety management;
Use of tools;
Cargo hold, engine room, boiler room, pump room.

B. *Safety requirements for ship entry and exit*

Information exchange;
Ship berthing, mooring and berthing requirements.

C. *General safety management requirements for ships in Port*

Obey the safety precautions and emergency procedures for terminal operations;
Relying on the management of mooring lines during the berthing period;
Fire monitoring;
Information and communication management;
Management of the crew;
Electrical equipment management.

D. *Safety requirements for ship loading and unloading in port*

Preparation measures before loading and unloading;
Safety measures during loading and unloading;
Handling requirements for cargo and ballast water;
Clearing and degassing operations;
Operating requirements for inert gas systems in fixed equipment;
Closed work requirements for the workplace.

E. *Emergency procedures*

Emergency plan and emergency handling procedures;
Emergency organization.

3 Comparison of Safety Management Indicators for Terminal Operations

A. *Safety requirements for loading, unloading, storage and transportation of bulk dangerous goods terminals in bulk in China*

(1) "Basic Requirements for Safety Technology of Oil Terminals" (GB16994-1997) [3]

"The Basic Requirements for Safety Technical Requirements for Oil Terminals" stipulates the basic safety requirements for oil terminal facilities, equipment and operations. Mainly includes:

- Location and layout of the terminal;

- Selection and setting of electrical equipment and power distribution;
- Anti-static, anti-stray current, lightning protection and other safety requirements;
- Firefighting facilities;
- Safety signs and warning signs;
- Process design requirements;
- Oil and heat pipeline design requirements;
- Railway loading and unloading oil facility design requirements;
- Design requirements for oil pump room and oil tank area;
- Anti-pollution facilities and equipment design requirements;
- Ventilation design requirements.

(2) "Technical Requirements for Port Handling of Bulk Liquid Chemical Products" (GB/T15626-1995) [4]

"The Technical Requirements for Port Handling of Bulk Liquid Chemical Products" stipulates the technical requirements for the process, equipment and facilities of bulk liquid chemicals in the port loading and unloading and transshipment process. Mainly includes:

- Handling process requirements;
- Flow rate requirements;
- Pipeline cleaning requirements;
- Facility requirements.

Configuration and use requirements for loading and unloading equipment such as pumps, hoses, oil delivery arms, storage tanks, and pipelines.

(3) "General Rules for Port Storage of Bulk Petroleum and Liquid Chemical Products" (GB17379-1998) [5]

"The General Rules for Port Storage of Bulk Petroleum and Liquid Chemical Products" stipulates the basic requirements for the storage of bulk petroleum and bulk liquidated products stored in storage tanks at ports. Mainly includes:

- Storage site requirements;
- Storage arrangements for storage tanks;
- Maintenance of goods and management of tank areas;
- Inbound and outbound management;
- Lightning protection, anti-static and fire-fighting measures;
- Pollution control and emergency measures;
- Staff training.

B. *International safety requirements for handling, storage and transportation of bulk liquid dangerous goods terminals*

Internationally, port operators are required to verify the relevant documents and certificates for the loading and unloading of dangerous goods in bulk, and confirm that the ship should have the ability to handle dangerous goods related to bulk liquids, understand and master the types, characteristics, operating procedures and emergency

measures of dangerous goods. Information and facilities for oil and gas recovery equipment. Specifically includes the following:

(1) Safety warning sign

Before loading and unloading dangerous goods in bulk, the port operator shall have appropriate safety warning signs at the entrances, exits, and frontiers of the terminal, and the signs shall be appropriate.

(2) Compatibility

The port operator shall ensure that the bulk dangerous goods in bulk are handled in good condition and that no dangerous reactions occur with other goods or materials.

(3) Communication

The port operator shall ensure that the communication during the loading and unloading process of bulk liquid dangerous goods is unblocked and equipped with explosion-proof communication equipment.

(4) Preparation measures before loading and unloading

- Prior to loading and unloading operations, the master and port operator should check the reliability of their cargo handling control systems, metering systems, emergency cut-off and alarm systems, and ensure that necessary safety equipment and safety clothing have been used.
- The captain and the port operator confirm the loading and unloading procedures in writing, including: maximum speed of loading and unloading, arrangement and capacity of ship-side cargo pipelines, maximum allowable pressure, arrangement and capacity of steam evacuation system, and possible pressure in emergency shutdown procedures. The possible accumulation of electrostatic charge, on-site workers during loading and unloading operations, and emergency measures and signals taken during loading and unloading operations.
- Complete and sign the safety checklist, which shall state the main safety measures before and during loading and unloading operations.
- The port operator should ensure that the start button of the material handler pump is off or within the control of authorized personnel.
- The port operator should ensure that the pipeline, the oil delivery arm and the hose are covered or covered with blind flanges when they are not suitable or in standby condition.

(5) Pipeline requirements

- The port operator shall ensure that the pipelines or movable pipelines provided on the terminal meet the following requirements:
- Install a suitable fixed line depending on the temperature and compatibility of the cargo.
- Take anti-collision protection measures to prevent damage.

- When handling flammable or flammable liquid bulk cargo, continuous conductivity should be provided unless an insulating flange or non-conductive hose is provided. The insulating member on the water side of the pipeline shall be continuously conductive with the hull, and the insulating member on the shore side of the pipeline shall be continuously conductive with the terminal grounding device.
- Insulating parts such as insulating flanges should be tested to ensure that no short circuit occurs. The terminal grounding device should be tested to ensure its effectiveness.
- Other metal connections between ships and shores should be protected to prevent sparks in flammable or flammable working environments.
- Safety management requirements for mobile lines (hose):
- The port operator should select the appropriate movable pipe according to the temperature, compatibility and working pressure of the cargo.
- The movable pipe should be prototype tested and hydraulically tested before use, and should be inspected regularly by visual inspection.
- The model, maximum working pressure and production date should be permanently and clearly marked on the movable pipe.
- The movable pipe should be equipped with an insulating flange.
- On-site monitoring should be arranged when using mobile pipe work.
- The port operator shall formulate relevant emergency measures for environmental protection, personnel and equipment protection in response to accidents that may occur during the removal process of the movable pipeline.
- After the movable pipe is used, the material should be emptied and cleared. If the material in the pipeline cannot be removed, measures shall be taken at both ends of the pipeline to prevent material gas from overflowing or air ingress.

(6) Oil delivery arm safety management requirements

- The port operator shall formulate emergency handling procedures for the operation, monitoring and disconnection of the delivery arm to protect the environment, personnel and equipment.
- The port operator should select the appropriate oil delivery arm according to the temperature, compatibility, working pressure or flow rate of the loading and unloading materials.
- The port operator shall take measures to ensure that the fuel delivery arm can empty the materials inside and outside the oil delivery arm before an emergency situation occurs after the oil delivery arm is disconnected from the pipeline.
- The working range of the oil delivery arm should be compatible with the ship.
- When multiple oil delivery arms are connected at the same time, there should be sufficient space.
- Port operators should regularly check the delivery arm to ensure that it is in good working order.

- The oil delivery arm should be equipped with an insulating flange.

(7) Material handling

- During the material transportation, the port operator and the master should take the following measures:
- Check frequently to ensure that the operation does not exceed the agreed back pressure and loading and unloading flow rates.
- Arrange the operators to observe that there are no leaks in the pipeline, the oil delivery arm, the movable pipeline and related equipment.
- Ensure the effectiveness of communication between ships and shores.
- Carry out safety inspections in accordance with the ship's shore safety checklist.
- According to the ship's cargo space, the loading operation is arranged to prevent material spillage.
- Arrange on-site monitoring of the operators.
- Ensure that the necessary safety equipment and safety clothing have been used.

(8) Prevention of leakage measures

Port operators should ensure that all drain lines are closed before loading and unloading operations and during loading and unloading operations to prevent water pollution when bulk liquid cargo leaks. In the event of a leakage accident, the port operator shall take effective measures for the collection and disposal of pollutants.

(9) Shore power supply

Port operators should ensure that shore communication cables that serve ships are safe for use in hazardous areas. The port operator shall not provide shore power services to the ship unless the shore power for the ship can be used safely in a flammable environment or in an emergency allowed by the port management department. When a ship carrying flammable or combustible cargo is loaded and unloaded by port, the shore power can only be used in the vicinity of the ported ship to ensure safety.

(10) End of work

- The port operator and the ship operator shall, within the scope of their respective duties, ensure that the loading and unloading valves shall be closed after completion of the loading and unloading operations, and the residual pressure in the pipelines, oil delivery arms and movable pipes shall be released unless the valves and terminals are required to be in operation. Open state.
- Before disconnecting the pipeline, the oil delivery arm, and the movable pipeline, the internal materials should be drained and the pressure released.
- The ship's shore connection line is covered with a blind plate.
- Ensure that the necessary safety equipment and safety clothing have been used.

C. *Comparative analysis*

(1) For the loading, unloading, storage and transportation of bulk liquid dangerous goods ports, taking into account the continuity and coordination of bulk liquid dangerous goods vessels and shore loading and unloading, loading and unloading in Port and terminal loading and unloading and storage of ships are generally regarded as two parts of the whole internationally. In China, the port administrative department is responsible for the safety supervision of the terminal operations. The safety supervision of the navigation and operation of the ship is the responsibility of the maritime department. The loading and unloading operations and dock loading and unloading operations of the ship during the port are applicable to different standard specifications, respectively, but due to different specifications. The drafting units are different, there are differences and inconsistencies between the ship and port loading and unloading operations.

(2) Since the international loading and unloading operations of bulk dangerous goods ships are often regarded as a link in the ship transportation process, relevant international conventions and guidelines have relatively complete and uniform provisions; Therefore, China is a party to the relevant conventions in bulk liquids. The safety management requirements for loading and unloading operations of dangerous goods ships in Port are basically the same as those in the international arena.

(3) Regarding the safety management of port terminal loading, unloading, storage and transportation operations, China and the international are basically consistent with the safety management principles of fire prevention, explosion protection and anti-static. Compared with the international, there is a lack of system regulations for the whole process of loading and unloading, before loading and unloading, after loading and unloading, etc., and lack of specific regulations on ship-shore linkage, the use of movable pipelines (hose), and shore power supply.

4 Suggestions

Further supplement and improve the standard specifications for port operations of bulk liquid dangerous goods. In the dock loading and unloading operations, the relevant standards are supplemented with relevant aspects such as ship-to-shore linkage, use of movable pipelines, and shore power supply, and attention should be paid to the coordination and unification of ship and shore loading and unloading operations.

Acknowledgements The authors would like to thank professor Xu Hongwei from the China Waterborne Transport Research Institute for his advising and help in this work. This title is funded by National Key R&D Program of China: Research on Key Technologies and Equipment of Transportation Infrastructure Construction Safety (Fund number: 2017YFC0805300).

References

1. Ministry of Transport of the People's Republic of China (2001) Safety operation rules for oil tankers and terminals, GB18434-2001. Standards Press of China, Beijing
2. International Maritime Organization (2006) International Safety Guide for Oil Tankers and Terminals: ISGOTT. Witherby Seamanship International Press
3. General Administration of Quality Supervision (1997) Inspection and Quarantine of the People's Republic of China, Basic Requirements for Safety Technology of Oil Terminals, GB16994-1997. Standards Press of China, Beijing
4. General Administration of Quality Supervision (1995) Inspection and Quarantine of the People's Republic of China, Technical Requirements for Port Handling of Bulk Liquid Chemical Products, GB/T15626-1995. Standards Press of China, Beijing
5. General Administration of Quality Supervision (1998) Inspection and quarantine of the People's Republic of China, General Rules for Port Storage of Bulk Petroleum and Liquid Chemical Products, GB17379-1998. Standards Press of China, Beijing

Research on the Influence of Traffic Factors Based in Mixed Logit Model on Short-Rent Lease of Housing

Bo Sun, Pengpeng Jiao and Yujia Zhang

Abstract With the aim to explore the influence of traffic factors on short-rent housing lease, this paper utilizes the utility function from the aspects of traffic convenience, traffic accessibility, generalized transportation cost, location advantage, road safety and traffic environment. We construct a mixed logit selection optimization model based on dis-aggregate theory, in which the reliability analysis and correlation analysis of SPSS are used as the basis for the tenant to decide short-rent housing. Based on the obtained questionnaire data of D-optimal method and the influence of different traffic modes and travel purposes on utility coefficient, maximum likelihood estimate (MLE) is introduced to solve the weight ratio, Matlab programming is used to solve to MLE, which shows that: expanding the traffic convenience and the accessibility of transportation infrastructure along the short rental housing source is the most effective means to stimulate the short rental market. Moreover, the reduction of transportation cost and the improvement of location advantage have a positive impact on short-rent housing rental satisfaction, demand and flexible price advantages. The evaluation value of road safety and traffic environment weight is on the low side, and the expected utility space is huge. Strengthening the service level of the vehicle will be conducive to share the running pressure of the rail transit.

Keywords Urban traffic · Short-rent housing lease · Utility function · Mixed logit model · SPSS · D-optimal method · Maximum likelihood estimate (MLE)

B. Sun · Y. Zhang
Beijing Advanced Innovation Center for Future Urban Design, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

P. Jiao (✉)
Beijing Urban Transportation Infrastructure Engineering Technology Research Center, Beijing University of Civil Engineering and Architecture, Beijing 100044, China
e-mail: jiaopengpeng@bucea.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_74

1 Introduction

At present, short-rent housing lease has become an important component of regional core settlements. In recent years, the use of the Internet to manage short-rent housing lease platforms emerge one after another. The germination of the concept of short-rent housing lease encourages the residents to use the shared housing resource in an efficient manner that improves the benefit of the idle house and promotes the maximization of the social benefit and cost. There is no doubt that the transportation industry reflects the development of economy and society. Therefore, there is a inseparable relationship between shortening the space to speed up urban construction and the implement of short-rent housing lease. Based on the overall consideration of urban transportation, to research the impact of traffic on short-rent housing lease through the comparison of the correlation between traffic network and short-rent leasing. Traffic factors stimulate the potential land use area and value of surrounding houses and provide a better development prospect for the short-rent housing lease market along the route.

Based on the principle of maximum utility and stochastic utility, dis-aggregate model combines the characteristics of traveler and traffic mode and it is widely used in the field of transportation for the time being. The mixed logit model is optimized by the traditional model. Mcfadden and Train [1] proposed and proved that the mixed logit model can simulate any utility model under the premise of selecting the appropriate mixed distribution function. It can be used to study the choice of short-rent house when the tenant considers the traffic factor. The current research results of domestic and foreign scholars in this field are as follows: Hensher [2] proposed the authenticity and use of the data extraction behavior required by the mixed logit model and the problems to be paid attention to in the practice of discrete selection method. Wang et al. [3] proposed the properties and estimation methods of mixed logit model, compared the two models, and proposed a mixed logit model estimation algorithm, which shows that the mixed logit model is more meaningful for estimating individual behavior in traffic mode selection. Luo et al. [4] established Logit model with different constraint conditions and two-distributed mixed logit model (MMNLl and MMNL2) to simulate traffic diversion problem. It is proved that the mixed Logit model can reasonably describe the unified preference factors without IIA restriction. It is more suitable for the model of traffic pattern division. Milton et al. [5] established a stochastic parameter model, using mixed Logit model to explain the road characteristics, environmental factors and driving behavior of highway sections, which can be used as an effective tool for highway safety planning. Bhat Chandra [6] came up with a quasi-random sequence method for Logit model estimation of mixed polynomials, which provides better planning and less time calculation in the case of intercity travel mode selection. Martínez et al. [7] put forward a mixed strategy that combines the effective compensation strategy within the selection domain with the cutoff factor that restricts the selection of the edge of the domain. Yáñez et al. [8] propose an approach to forecasting attitudes and perceptions which can influence users' behaviour, their results show that mixed models are clearly superior to even

highly flexible traditional models. Huamin et al. [9] established a mixed Logit model for traffic mode selection, which allows coefficients to be combined with different distribution types to simulate random number sequences and calculate simulation probability, which clearly explains the various factors that affect people's travel. Hu et al. [10] use mixed Logit model to calculate the sharing rate of driving distance, time, speed and cost, and evaluate the traffic sharing rate of underground expressway, which is more reasonable than the traditional Logit model. Romo et al. [11] studied the influence factors of multi-lane interstate road collision in the same direction (that is back end, angle, side slide), established mixed logit (MXL) model, estimated back end, the angle and the probability of side-slip collision act as a function of the vehicle following attribute and other driving maneuvers before the collision. Haleem and Gan [12] used mixed logit model estimation to explore models that are difficult to quantify and may have an impact on unobserved accidents, and to determine the geometry, traffic, and environment of the severity of collision damage on urban expressways. Vehicle correlation and driving Staff correlation predictors. Srikukenthiran et al. [13] explore the species of Toronto bus station where pedestrians' choice between common staircases and escalators is modeled using a standard set of binary and mixed logit models, its prediction performance is always better than that of standard model, showing good prediction ability (nearly 90%). Yang et al. [14] estimated the travel time value of urban rail transit by using Mixed Logit model of uniform distribution, normal distribution and logarithmic normal distribution, and accurately predict the impact of urban rail transit demand management policy on travel behavior. Lovreglio et al. [15] use an online stated preference survey that carried out making use of non-immersive virtual reality and a mixed logit model is calibrated using these data. Their findings can assist in designing and managing building and transportation systems. Tang et al. [16] constructed a calibration model of observation data based on mixed Logit selection. The significance and symbol of the parameters explained the key factors of the choice of travel mode in the context of urban comprehensive traffic and analyze the marginal effect of taxi and net fare.

With the increase of computer speed and the emergence of simulation algorithms, the mixed logit model with rich physical meaning has been obtained more and more application. Mixed logit model is a highly adaptive model that overcomes two important defects of traditional logit model: random preference restriction and IIA requirement, which can intuitively express personal preference information. This paper discusses the priority and feasibility of different traffic conditions for different tenants to choose short rental housing from six aspects of traffic factors, and establishes utility function through dis-aggregate theory and stochastic utility theory. The maximum likelihood estimation (MLE) was used to process the questionnaire data and calculate the utility function parameters. Besides, this paper introduce the mixed logit model to do further research.

2 The Characteristic Analysis of Tenant Choosing Short-Rent Housing Lease

With the acceleration of urbanization, shared culture has become the mainstream of social development. As one of the shared cultures, short-rent housing lease brings lots of social and economic benefits. There is no doubt that transportation has a direct impact on the choice of short-rent rental housing. Moreover, different factors in the transportation industry generates different effects on the benefits of short-rent housing. Whereas the vigorous development of the short-rent market that can reflect the advantages and disadvantages of regional traffic to a certain extent. At the same time, traffic factor is an important reference for the choice of short rental housing based on the needs of medical treatment, family visit, travel and education. There are six aspects as follows: traffic convenience, traffic accessibility, generalized transportation cost, location advantage, road safety and traffic environment.

The improvement of the traffic conditions can increase the comprehensiveness of the nature of the land, thereby affecting the location of social activities, stimulating new land development and advancing social and economic growth. Through the transportation allocation and land use, the circulation of land use and traffic system has been launched until it tends to balance. The leaseholder sets the traffic impedance through the traffic chain, the choice probability of short-rent house lease fluctuates up and down with the change of six different traffic factors. According to different leaseholders' traffic modes and travel purposes, the degree of preference and the basis for consideration of traffic conditions are also different to some extent with the various selection probability (Fig. 1).

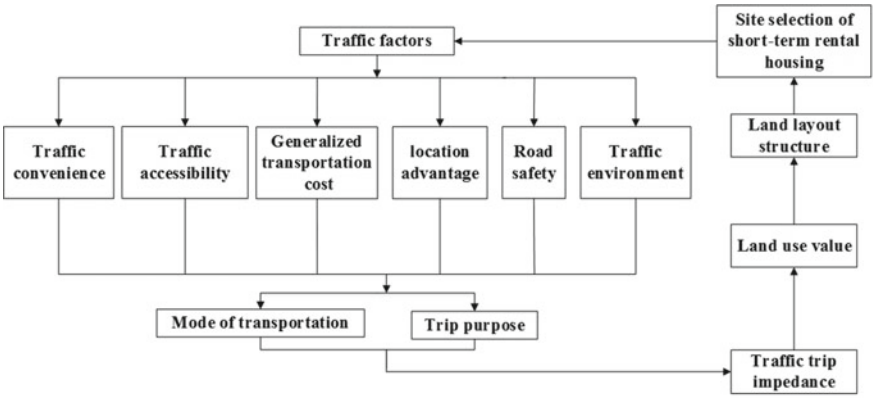


Fig. 1 Characteristic analysis chart of short-rent housing selected by leasors

3 Establishment of Utility Function

According to the basic theory of utility, the tenant always chooses the most comprehensive traffic utility to choose the short rental house. The utility value is composed of two parts: fixed term and random item. Assuming that the renter k may choose a set of short-rent rentals as D_n, in which the combined traffic utility of house j is U_{kj}, the tenant chooses the following terms for short rental i:

$$U_{ki} > U_{kj} \tag{1}$$

From Eq. (1), $i \neq j$, $j \in D_n$.

Based on the utility maximization theory, renters in face of different traffic factors, the selection probability P_{ki} of short rental housing i as (2).

$$P_{ki} = \int (U_{ki} > U_{kj}) f(\phi|\theta) d\phi = \int (W_{ki} + \mu_{ki} > W_{kj} + \mu_{kj}) f(\phi|\theta) d\phi \tag{2}$$

From Eq. (2), $0 \leq P_{ki} \leq 1$, $\sum_{i=D_n} P_{ki} = 1$, W_{ki} is the utility function fixed item of lessee k for short-rent rental house i according to different transportation factors; μ_{ki} is the utility function of lessee k that selects i for short-rent rental according to different transportation factors.

If $W_{ki} = \delta_{ki} X_{ki}$; μ_{ki} follows the distribution of double exponential, the distribution function of each probability term $\mu_{k2}, \ldots, \mu_{kj}$ is $F(\mu_{k1}, \mu_{k2}, L, \mu_{kj})$, and the probability density function is $f(\mu_{k1}, \mu_{k2}, L, \mu_{kj})$.

4 Utility Function of Different Traffic Factors

The fixed items in utility value are mainly measured in line with the economic attributes, travel characteristics and travel mode characteristics of the renters. Referring to the comprehensive consideration of utility function traffic factors, this paper discuss the six indexes of traffic convenience, traffic accessibility, generalized transportation cost, location advantage, road safety and traffic environment.

The convenience of traffic mainly depends on the convenience of different traffic modes from the short rental house to the destination, expressed by B_k and measured by time dimension, as (3).

$$B_k = (T_i + t_{gi} + t_{di} + t_{qi}) l(t) \cdot z(t) \tag{3}$$

From Eq. (3), T_i refers to the transfer time between the i short rental house and the destination and different modes of transportation; t_{gi} is the purchase time between the destination and the i short rental house; t_{di} is the waiting time between the destination and the i short rental house; t_{qi} refers to the other time spent between the destination

and the i short rental house; $l(t)$ is the travel efficiency of the renter; $z(t)$ is the travel time value of the renter.

Transport accessibility indicates the degree of difficulty in using or approaching (traffic) facilities around the tenant's predefined short rental housing expressed in L_k, as shown in (4).

$$L_k = \ln \sum_{i=1}^{m_{ki}} \exp(V_i) \tag{4}$$

From Eq. (4), m_{ki} is the choice of the tenant's travel mode, V_i is the choice utility of the travel mode.

The broad sense of transport costs takes into account the transportation costs that a tenant spends on a variety of means of transport, including parking charges, subway, public transport, shared bicycles, and train, high-speed rail, and air travel costs, expressed in C_k. The main consideration here is cost and congestion cost.

$$C_k = q_k + y_k \tag{5}$$

From Eq. (5), q_k represents a collection of various transportation costs; y_k represents the cost of congestion caused by taking a vehicle.

The traffic around the short rental housing source covers the commercial area, park, supermarket, scenic area and so on. It increases the traffic flow, logistics, people flow of residential area, and also promotes the commercial development and the promotion of business atmosphere in the vicinity of the line, as indicated by Q_k, as shown in (6).

$$Q_k = \frac{\sum_{i=1}^{n} m_i + o_i + \sum_{i=1}^{n} p_i}{3} \tag{6}$$

From Eq. (6), m_i indicates the quality scale around the borderline of the i short rental house, the average road accessibility around the i short lease house, and p_i represents the land use level around the i short lease house.

Road safety is expressed as the safety coefficient of the road around the short-rent rented house. With the increase of urban motor vehicles and speed, the number of accidents, injuries, deaths and the direct economic losses increase significantly over the past few years, so houses with more complete road infrastructure are more favored by renters. It is particularly important to relieve traffic stress and eliminate fatigue, as indicated by E_k.

$$E_k = \frac{t\max}{\alpha_i \exp(-\beta_i t) + 1} \tag{7}$$

From Eq. (7), $t_{\max}$ represents the maximum time to recover from fatigue; t is the time taken by means of transportation; α_i is the dimensionless coefficient; β_i is the strength coefficient of fatigue recovery in a fixed time.

The traffic environment is expressed as the core living conditions of the whole environment of the short rental housing area. Such as noise pollution, tail gas pollution, light pollution and so on, have caused damage to the living environment of tenants, environmental factors directly affect the market value of short-rent lease that are of great significance to the evaluation of short-rent rental market, as indicated by H_k, as shown in (8).

$$H_k = \sqrt{\sum_{i}^{n} (h_i \times q_i)^2} \tag{8}$$

From Eq. (8), h_i denotes the influence factor on traffic environment and q_i indicates the weight value of the influence factor of item i.

Based on the above six indexes and the linear function relation of fixed utility in utility function, the global utility function expression U_{ki} is established as shown in (9).

$$U_{ki} = W_{ki} + \mu_{ki} = \delta_1 B_k + \delta_2 L_k + \delta_3 C_k + \delta_4 Q_k + \delta_5 E_k + \delta_6 H_k + \mu_{ki} \tag{9}$$

From Eq. (9), $\delta_1 \sim \delta_6$ is the coefficient of utility function fixed term.

5 Modeling and Coefficient Solving

Modeling

According to the different traffic factors of Disaggregate Model, to research the influence of short rental housing selection mode. Due to the traffic factors are not independent of each other, in order to avoid the IIA characteristics of the MNL model and make up for the shortcomings of the traditional model, the choice rate P_{ki} of the short rental house based on the traffic factor is in accordance with the mixed logit model. The expression is,

$$P_{ki} = \int \left[\frac{\exp(\alpha^T \times U_{ki})}{\sum_{i \in D_k} \exp(\alpha^T \times U_{ki})} \right] f(\phi|\theta) d\phi \tag{10}$$

From Eq. (10), D_k is a set of short rental houses that tenants can choose; α^{T} is a random vector to be estimated; vector ϕ is a random variable of traffic elements; and θ is a characteristic parameter.

The selection probability of the mixed Logit model can be regarded as the weighted average of the selection probability of the multidimensional Logit model, and the weight is determined by the distribution density function $f(\phi|\theta)$. Vector α can be divided into normal distribution, logarithmic distribution, exponential distribution and uniform distribution. Moreover, different distribution functions have their own

merits and demerits, which need to be judged according to the research content and practical experience, and comprehensive data acquisition.

The utility value is measured by referring to the characteristics of integrated traffic mode in the study of utility function factors, according to the utility linear function relation of the above six indexes and utility parameters. A mixed logit model of tenant selection value and short rental house selection probability based on traffic influence factors is established as follows,

$$P_{ki} = \int \left\{ \frac{\exp[\alpha^T \times (\delta_1 B_k + \delta_2 L_k + \delta_3 C_k + \delta_4 Q_k + \delta_5 E_k + \delta_6 H_k + \mu_{ki})]}{\sum_{i \in D_k} \exp[\alpha^T \times (\delta_1 B_k + \delta_2 L_k + \delta_3 C_k + \delta_4 Q_k + \delta_5 E_k + \delta_6 H_k + \mu_{ki})]} \right\} f(\phi|\theta) d\phi \tag{11}$$

Coefficient solving

The probability selection formula of mixed Logit model is not closed, so the analytical method can not be used to obtain the results directly. Using the method of simulation to calculate the probability and determine the best weight ratio, Maximum Likelihood Estimate is used in this paper. On the premise of θ, Halton is used to extract random vector ϕ in density function $f(\phi|\theta)$, which is denoted as ϕ^k. In the first extraction, the $k = 1$ is recorded, the $L_{ki}(\phi^k)$ is calculated according to Formula (12), and when K times (500–1000 times) are repeated, the average value of the $L_{ki}(\phi^k)$ calculation results is taken as the probability simulation value, that is, $\widehat{P_{ki}} = \frac{1}{K} \sum_{k=1}^{K} L_{ki}(\phi^k)$.

$$L_{ki}(\phi) = \frac{e^{W_{ki}}}{\sum_j e^{W_{kj}}} \tag{12}$$

The parameters of utility function are solved by Maximum Likelihood Estimate. If the tenant k has a fixed utility $W_{ki} = \delta_{ki} X_{ki}$, and then $\delta_{ki} = (\delta_{1ki}, \delta_{2ki}, \delta_{3ki}, \delta_{4ki}, \delta_{5ki}, \delta_{6ki})$, $X_{ki} = (B_{ki}, L_{ki}, C_{ki}, Q_{ki}, E_{ki}, H_{ki})^{\mathrm{T}}$, δ_{ik} is a parameter vector. If the sample number of renters is N, it is defined as:

$$\gamma_{ki} = \begin{cases} 1, k \text{ choose } i \\ 0, \text{else} \end{cases} \tag{13}$$

The likelihood function L^* is shown in (14):

$$L^*(\delta) = \prod_{k \in N} \prod_{i \in D_k} \widehat{P_{ki}^{\gamma_{ki}}}(i|M; \delta)(k = 1, 2, L, N; i \in D_k) \tag{14}$$

Take the logarithm of both sides of Formula (14) and combine with W_{ki} to obtain the likelihood function of logarithm, as shown in Eq. (15):

$$\ln L^*(\delta) = \sum_{k=1}^{N} \sum_{i \in D_n} \gamma_{ki} \ln \widehat{P_{ki}}(i|M;\delta) \tag{15}$$

From Eq. (15), M is the collection of six traffic factors for the renter to select the short-rent rental house, as shown in (16):

$$P_{ki} = \frac{\exp[\eta\delta_{ki}X_{ki}/(\delta_{1k}X_{1k} + \delta_{2k}X_{2k} + \delta_{3k}X_{3k} + \delta_{4k}X_{4k} + \delta_{5k}X_{5k} + \delta_{6k}X_{6k})]}{\sum_{j=1}^{6} \exp[\eta\delta_{ki}X_{ki}/(\delta_{1k}X_{1k} + \delta_{2k}X_{2k} + \delta_{3k}X_{3k} + \delta_{4k}X_{4k} + \delta_{5k}X_{5k} + \delta_{6k}X_{6k})]} \tag{16}$$

According to the nature of maximum likelihood estimate, to change the value of θ and maximum likelihood estimate is simulated to obtain the maximum value of Formula (15). Partial derivative is assumed to be $\partial L^*(\delta)/\partial\delta = 0$ and the likelihood function is integrated. Iterative algorithm is used to solve the weight coefficient of each traffic indicator. After matching the corresponding utility value, the weight ratio of the six traffic factors is calculated (among them, there are K tenants who consider the traffic factor of item i).

6 Case Analysis

Questionnaire design and survey plan

The questionnaire is divided into three parts: the first part, the socio-economic attribute information of the short-term house leasing selector; the second part is the RP questionnaire, which is used to understand the frequency, reason and travel distance of the different modes of travel of the renter; the third part is the SP questionnaire, in the context of commuting and non-commuting, the attributes in the model are combined into different situations according to different levels. The measurement methods of each variable are shown in Table 1.

In order to obtain more reliable parameter estimation, the D-optimal design method was used to generate the declarative preference questionnaire, and the Matlab software programming was used to design the whole factor design method. From this, 60 test combinations were randomly selected, D_{error} was calculated, and iteration was continuously performed to find the 60 combined tests with the smallest D_{error}. In order to avoid the confusion of the respondents or the random confusion caused by over-reporting too many situations, 6 kinds of questionnaires were randomly selected to form one questionnaire, so there were 10 kinds of questionnaires.

Beijing was chosen as the investigation city and short-rent platforms as the research object. Finally, Mayi, Xiaozhu, Airbnb, Tujia, Muniao are chosen as the research object. These short-rent platforms are representative in the user experience and evaluation, the number of houses, the coverage area and so on. The survey used age-based stratified random sampling method to conduct questionnaire surveys in different time periods. The total number of valid samples collected in this study was 513. Users of different age groups and different age groups could participate in

Table 1 Variable design

Attribute	Variable	Code	Survey method
Socioeconomic attribute	Gender	X_{Gender}	Men: 1, Women: 0
	Age	X_{Age}	<18 years: −5, [18, 26) years: −3, [26, 36) years: −1, [36, 46) years: 1, [46, 60) years: 3, ≥60 years: 5
	Income	X_{Income}	≤3000 yuan: −3, (3000, 4500] yuan: −1, (4500, 6000] yuan: 0, (6000, 8000] yuan: 1, >8000 yuan: 3
Travel mode attribute	Travel purpose	X_{MD}	Leisure and entertainment: 1, Non-leisure entertainment: 0
	Connection time	X_{AM}	Walk/min: 10, 15, 20, 25 Non-motor vehicle/min: 8, 10, 12, 15 Conventional bus/min: 5, 7, 10, 15 Vehicle/min: 3, 5, 7, 10 Rail transit/min: 2, 4, 6, 8
	At the car time	X_{BM}	Walk/min: 10, 15, 20, 25 Non-motor vehicle/min: 8, 10, 12, 15 Conventional bus/min: 5, 7, 10, 15 Vehicle/min: 3, 5, 7, 10 Rail transit/min: 2, 4, 6, 8
	Parking time	X_{SM}	Walk/min: 10, 15, 20, 25 Non-motor vehicle/min: 8, 10, 12, 15 Conventional bus/min: 5, 7, 10, 15 Vehicle/min: 3, 5, 7, 10 Rail transit/min: 2, 4, 6, 8
	Time of arrival	X_{DM}	Walk/min: 10, 15, 20, 25 Non-motor vehicle/min: 8, 10, 12, 15 Conventional bus/min: 5, 7, 10, 15 Vehicle/min: 3, 5, 7, 10 Rail transit/min: 2, 4, 6, 8
	Travel of distance	X_{TD}	5 km: −1, 10 km: 0, 15 km: 1
	Cost	X_{TC}	(−1 indicates low level, 0 indicates medium level, 1 indicates high level) Walk: 0 Non-motor vehicle (yuan/time): 0.5, 1 Conventional bus (yuan/time): 1, 2, 4 Vehicle (yuan/km): 1.2, 2.2,3.2 Rail transit (yuan/time): 2, 4, 7

the questionnaire. The formal survey is divided into two phases: the first phase, the electronic questionnaire is distributed through the questionnaire platform, WeChat, email, etc. In the second phase, based on the age distribution characteristics of the samples obtained in the first phase, the paper questionnaire is distributed in a targeted manner. Response rate and questionnaire are efficient, and each respondent who completes the questionnaire can get a cash reward of 20 yuan.

7 Analysis of Survey Data

Reliability analysis

Reliability analysis, also known as reliability analysis, is used to measure the degree of consistency of a research scale under different measurements. In this study, Cronbach α coefficient is used to measure the reliability of each variable. With the help of SPSS software, the results are shown in the Table 2.

From table, we can see that there are six variables involved in this study, and the Cronbach α values of these six variables are 0.892/0.881/0.913/0.854/0.866/0.879 respectively, all of them are above 0.8, the minimum is 0.847, and the maximum is 0.913, which indicates that the reliability of the variables involved in this study is very high. The sample answer is accurate and reliable, which means that the sample data in this study can be used for further study.

Correlation analysis

This part uses the correlation analysis method to study the correlation between the different traffic factors of the sample and the possible influence of the tenant to choose the house and so on. The correlation is expressed by Pearson correlation coefficient. The correlation coefficients between the evaluation results and the factors are, 0.711/0.732/0.725/0.538/0.704/0.622.

In particular, there are significant positive correlations between the choice of tenant and the six variables, such as traffic accessibility, transportation cost, location advantage, road safety, traffic environment, etc. The correlation coefficient with traffic accessibility is 0.732 and the correlation significance is the most obvious, which

Table 2 Reliability of different traffic factors

Variable name	Cronbach α
Traffic convenience	0.892
Traffic accessibility	0.881
Generalized transportation cost	0.913
Location advantage	0.847
Road safety	0.866
Traffic environment	0.879

means that the tenant is sensitive to the availability of traffic and may not improve the expectation of travel efficiency in the near future.

Weight calculation of short-rent housing lease selection influenced by traffic factors

A questionnaire survey will be conducted among those who use high speed rail, aircraft, public transport, automobile, subway, bicycle and other means of transportation as means of travel for the purpose of traveling, seeking medical treatment, school study, business work, visiting relatives and friends, and then to calculate the fixed utility function coefficients of the multidimensional mixed Logit model based on the survey data and the maximum likelihood estimation, and to further calculate the corresponding utility values according to the relevant data.

Influencing factors of utility function

Because the time value and travel efficiency of different income passengers are different, the corresponding data of different traffic factors are shown in the following Table 3.

Assuming that the working day is 8 h and working 22 days a month, according to the income of the different destination population, the corresponding time value table is calculated as follows (Tables 4 and 5).

Weight coefficient

For the fixed utility coefficients, combined with the survey results and the relevant data in Table 2, the software of Matlab is used to solved maximum likelihood

Table 3 Data of traffic factors under different traffic modes

Mode of transportation	Traffic factors					
	Traffic convenience	Traffic accessibility	Generalized transportation cost	Location advantage	Road safety	Traffic environment
Walk	43/60	0.99	0.00	0.63	0.51	0.20
Non-motor vehicle	50/60	0.62	0.5	0.61	0.33	0.24
Conventional bus	52/60	0.69	1.0	0.51	0.34	0.12
Vehicle	51/60	0.55	13.8	0.32	0.28	0.27
Rail transit	55/60	0.85	4.0	0.57	0.60	0.44

Table 4 Time value of leaseholders for different purposes

Purpose of choosing short rent housing	Travel	Hospitalize	Study	Business office	Visit relatives and friends
Time value (yuan/h)	28.89	13.35	25.33	37.31	16.79

Table 5 Travel efficiency of tenants with different purposes

Purpose of choosing short rent housing	Travel	Hospitalize	Study	Business office	Visit relatives and friends
Trip efficiency	0.72	0.91	0.85	0.83	0.68

estimation which can calculate the fixed utility function coefficients of the mixed Logit model according to Eqs. (14)–(16). The corresponding weight coefficients are obtained as shown in Table 6.

Fixed utility value

Based on the relevant data and Formula (9) of Tables 2, 3 and 4, the fixed utility value of short rental housing selected by tenants with different traffic modes for different purposes in the mixed Logit model is calculated under six traffic factors, such as Table 7.

Weight ratio calculation

The fixed utility value in Table 7 is added to the mixed Logit model to calculate the traffic modes of different renters, and the weight of six traffic factors, such as Table 8,

Table 6 Weight coefficient table for different travel purposes

Purpose of choosing short rent housing	Weight coefficient					
	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6
Travel	27.853	23.676	8.189	6.668	6.124	2.809
Hospitalize	31.172	28.332	5.002	12.335	−10.897	−51.333
Study	20.011	11.174	22.565	9.989	10.329	12.218
Business office	28.727	21.003	2.985	4.204	9.393	6.664
Visit relatives and friends	18.383	15.265	6.773	−0.630	1.871	−7.319

Table 7 Fixed utility value of different traffic modes

Mode of transportation	Traffic factors					
	Traffic convenience	Traffic accessibility	Generalized transportation cost	Location advantage	Road safety	Traffic environment
Walk	9.231	28.459	0.000	8.133	20.673	11.505
Non-motor vehicle	14.461	10.309	2.997	8.565	19.726	11.218
Conventional bus	12.208	9.569	3.332	7.003	21.212	13.996
Vehicle	18.448	19.238	15.538	6.115	23.665	12.653
Rail transit	20.393	13.122	4.991	5.381	35.288	19.355

Table 8 Weight proportion of six traffic factors

Mode of transportation	Traffic factors					
	Traffic convenience	Traffic accessibility	Generalized transportation cost	Location advantage	Road safety	Traffic environment
Walk	0.535	0.842	0.000	0.234	0.087	0.103
Non-motor vehicle	0.644	0.611	0.387	0.379	0.101	0.088
Conventional bus	0.633	0.647	0.422	0.384	0.117	0.091
Vehicle	0.731	0.692	0.581	0.217	0.093	0.105
Rail transit	0.705	0.649	0.457	0.345	0.136	0.129

is given.

As can be seen from Table 8 and Fig. 2, in the process of choosing a short rental house, the traffic factors considered by the tenant have obvious differentiation and priority. Among them, traffic convenience and accessibility are the most important traffic factors for the tenant, no matter what kind of purpose or mode of transportation they use, the weight is very high, and the traffic accessibility is the most important factor for the tenant to take into account. It means that the utility brought about by the two is also the greatest. The second factor is the generalized transportation cost and location advantage. The regional attention paid to road safety and traffic environment is relatively weak and may even be neglected to some extent. There are potential risks of traffic hazards indicating that the renter has a low awareness of traffic safety and threats to the traffic environment.

Marginal effect analysis

The marginal effect analysis of the actual distance between the short rental house and the destination selected by the tenant based on different purposes, the direct marginal utility represents the influence of the change of each unit variable of a certain traffic mode of different destination on its selection probability, that is,

$$J_{M_{kj}}^{P_{kj}} = \frac{\partial P_{kj}}{\partial M_{kj}} \tag{17}$$

The cross marginal utility means the influence of the change of each unit variable of different destination on the corresponding transportation mode selection probability. That is,

$$J_{M_{ki}}^{P_{ki}} = \frac{\partial P_{ki}}{\partial M_{ki}} \tag{18}$$

In Eqs. (17), (18), M_{kj} and M_{ki} represent the variation of each unit variable of different destinations, and calculate the direct and cross marginal utility of each small

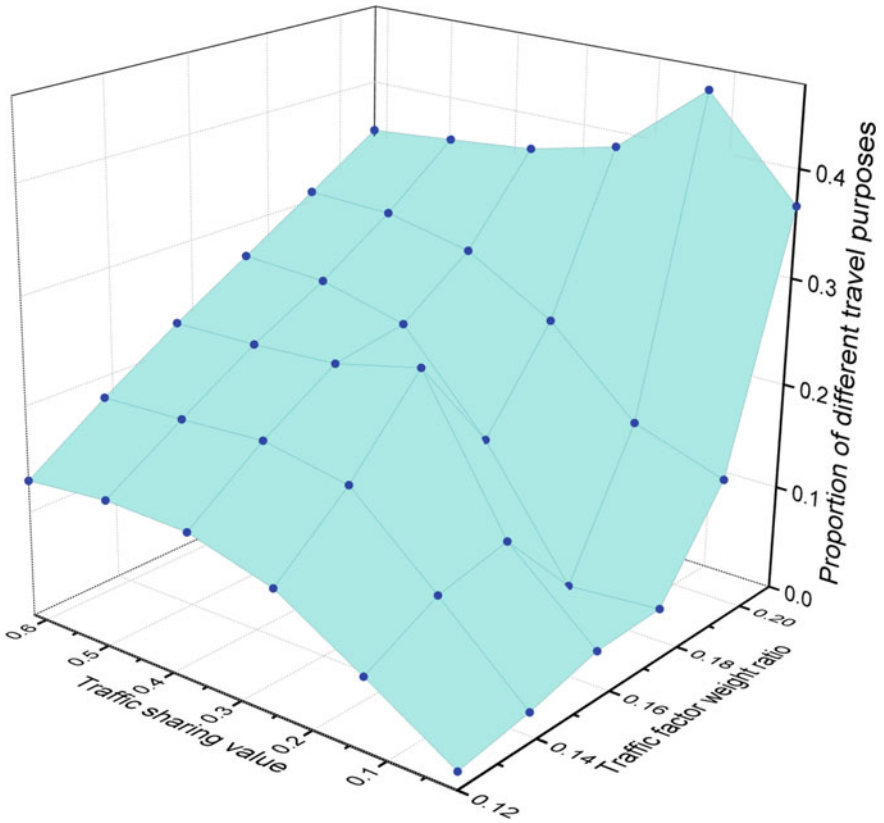


Fig. 2 The figure of different travel destinations, traffic sharing values and traffic factor weight conversion ratios

increase in the distance between the short rental house and the destination, as shown in the following table.

As seen in Table 9, for each 1 km increase in distance between a short rental house and a destination, the probability of walking and using non-motor vehicles to travel to the destination under normal commuting conditions decreased by 6.52% and 3.93% respectively. The probability of choosing conventional bus and automobile travel will increase by 1.36% and 0.23% respectively, while the probability of choosing rail transit commuting will increase 7.71%. It can be seen that with the small increase

Table 9 Distance marginal effect analysis table

Variable change value	Walk	Non-motor vehicle	Conventional bus	Vehicle	Rail transit
Each increase in distance by 1 km	−6.52	−3.93	+1.36	+0.23	+7.71

of distance rail transit attracts the tenants who choose walking and non-motor vehicles to play an obvious role in substitution and the substitution of automobile travel is also relatively limited for its unique advantages.

8 Conclusion

Based on the questionnaire survey results of Beijing and the data of short-rent platforms, with the help of the utility model of various traffic influencing factors, absolute utility is replaced by relative utility to avoid some indexes falling into local isolated changes. Moreover, according to the characteristics of stratification of six traffic factors, to construct a mixed logit model for optimizing the choice of short rental housing step by step. The following conclusion is obtained by using maximum likelihood estimation solution and marginal effect method.

1. The choice of short-rent rental houses is the result of the comprehensive effects of different travel purposes, travel modes, travel characteristics and different traffic factors on renters, and there is an obvious hierarchical relationship between the influences and the influences of various traffic factors. The mixed logit model fully reflects the hierarchical relationship among the variables and is more suitable for the study of traffic and short-rent housing rental problems. In this sense, it provides a theoretical basis for the validity of tenants' choice.
2. In view of the evaluation of various traffic factors in the model, most of the tenants take the more accessible traffic mode and less commuting time as the core purpose without comprehensive and systematic comprehensive indicators, it can be seen that the tenants' awareness of traffic safety is weak, and the road environmental are often ignored, resulting in increased distractions, potential traffic hazards and traffic conflicts.
3. Rail transit is the main way for tenants to replace walking and non-motor vehicle travel, which may lead to excessive operation pressure. Therefore, under the overall operating conditions of urban traffic, the substitution role of automobile travel is limited as well. Therefore, some measures need to be put into practice to encourage taxis and ride-sharing, share car rental and other car services. Besides, due to the lack of theoretical content at present, this paper has important significance and reference value for traffic management department and tenant of short-rent housing lease.

Declaration of conflicting interests The authors declare that there is no conflict of interests regarding the publication of this paper.

Funding This research has been supported by National Natural Science Foundation of China (51578040) and Beijing Natural Science Foundation (8162013). Valuable comments and suggestions from anonymous reviewers are also gratefully acknowledged.

References

1. Mc Fadden D, Train K (2000) Mixed MNL models for discrete response. J Appl Econ 15(5):447–470
2. Hensher DA, Greene WH (2003) The mixed logit model: the state of practice. Transportation 30(2):133–176
3. Wang S, Wei H, Lu Z (2006) Study on mixed logit model and its application in traffic mode split. J Highw Transp Res Dev 5:88–91
4. Luo J, Li X, Li X (2007) Application of mixed multinomial logit models in urban traffic split of Hefei City. In: International conference on wireless communications. IEEE
5. Milton JC, Shankar VN, Mannering FL (2008) Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accid Anal Prev 40(1):260–266
6. Bhat Chandra R (2008) Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transp Res Part B 35(7):677–693
7. Martínez F, Aguila F, Hurtubia R (2009) The constrained multinomial logit: a semi-compensatory choice model. Transp Res Part B Methodological 43(3):365–377
8. Yáñez MF, Raveau S, Ortúzar JD (2010) Inclusion of latent variables in mixed logit models: modelling and forecasting. Transp Res Part A Policy Pract 44(9):744–753
9. Huamin LI, Huang H, Liu J (2010) Parameter estimation of the mixed logit model and its application. J Transp Syst Eng Inf Technol 10(5):73–78
10. Hu Y, Ling M, Shang H (2011) Evaluation on effects of underground expressway based on mixed logit model. In: International conference of chinese transportation professionals
11. Romo A, Hernandez S, Cheu RL (2013) Identifying pre-crash factors between cars and trucks on interstate highways: mixed logit model approach. J Transp Eng 140(3):272–273
12. Haleem K, Gan A (2013) Effect of driver's age and side of impact on crash severity along urban freeways: a mixed logit approach. J Saf Res 46:67–76
13. Srikukenthiran S, Shalaby A, Morrow E (2014) Mixed logit model of vertical transport choice in Toronto subway stations and application within pedestrian simulation. Transp Res Procedia 2:624–629
14. Yang L et al (2015) Calculation of travel time value for urban rail transit based on mixed logit model. Urban Mass Transit 18(5):29–32
15. Lovreglio R, Fonzone A, dell'Olio L (2016) A mixed logit model for predicting exit choice during building evacuations. Transp Res Part A Policy Pract 92:59–75
16. Tang L et al (2018) Choice behavior of taxi-hailing based on mixed-logit model. J Transp Syst Eng Inf Technol 18:108–114

An Accurate Prediction Method for Airport Operational Situation Based on Hidden Markov Model

Xintai Zhang, Yanwen Xie, Yaping Zhang, Zhiwei Xing, Xiao Luo and Qian Luo

Abstract This paper is mainly devoted to an prediction method for airport operational situation which is one of the most important parts of the airport operation system. In order to provide theoretical support for high-level airport management, field operation management, air traffic control and airlines, and improve the service capacity of the airport, this paper makes a prediction study of the airport operation situation. Hidden Markov (HMM) prediction model is established based on the analysis of airport operation system. Baum-Welch and Viterbi algorithms are used to solve the prediction results. The model is validated and applied in a domestic hub airport. The results show that the prediction accuracy of HMM is 60 and 20% higher than that of Autoregressive Moving Average Model and Grey Markov model, respectively. It can also improve the situation value of airport operation situation, i.e. airport service capability. This method is more suitable for the analysis of airport operation.

Keywords Airport traffic · Operational situation forecast · Hierarchical division · Hidden markov model

1 Introduction

The development of civil aviation industry promotes the development of transportation industry and enriches the way people travel. The booming demand of air transportation leads to the shortage of service resources, which reduces the efficiency

X. Zhang · Y. Xie · Y. Zhang (✉)
School of Transportation Science and Engineering, Harbin Institute of Technology, Heilongjiang Province, Harbin 150000, China
e-mail: zxlt0905@163.com

Z. Xing
Research Base of Ground Support Equipment, Civil Aviation University of China, Tianjin 300000, China

X. Luo · Q. Luo
The Second Institute of Civil Aviation Administration of China, Chengdu, Sichuan 610000, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_75

and service capacity of the entire airport and weakens the advantages of air transport. To improve the operation efficiency and service capacity of the airport, many statistical model have been proposed and utilized, such as Hidden Markov Model, Autoregressive Moving Average Model, Gray Theory, Support Vector Machine and Neural Network, to explain the whole operation process of the airport and predict the problems that will occur and provide solutions. Therefore, it is necessary to study the airport operation situation.

The understanding of situation originates from foreign countries. Endsley defines situation awareness as 'In a certain space and time conditions, the acquisition of understanding of environmental factors and the prediction of the future state' [1]. The concept of situational awareness has been widely studied and applied in military battlefield, nuclear reaction control, force system and network security [2, 3]. The earliest application research of situational awareness in civil aviation mainly focuses on the field of air traffic management [4, 5]. In China, Xie Lixia and others studied the network situation and network security situation [6]. In 2014, Xu Xiaohao, Ren Jie, Li Nan and others identified and studied the traffic situation in the terminal area of the airport [7, 8]. In 2015, Wei Ran used the multi-time scale method to forecast the expressway operation situation [9]. In 2016, Wang Guangchao and Qian Kun respectively estimated the flight operation situation based on GIS and designed the traffic operation situation analysis system of intelligent city transportation system [10, 11]. In 2017, Yuan Liquan studied the dynamic traffic characteristics and operation situation of terminal area [12]. In 2017, Feng Chao and others analyzed the characteristics of data changes of air-to-air aircraft, proposed an analysis method of close-range Air-to-Air Combat based on Hidden Markov, predicted the flight state sequence, and finally obtained the air combat decision point [13]. In 2018, Xu Cheng and others used Bilevel Programming Model to study airport operation efficiency from taxiway and parking assignment [14]. Domestic and foreign research on operational situation is more about assessment and analysis. Referring to the relevant literature of network security situation prediction, we can see that the main methods of situation prediction are Autoregressive Moving Average Model ARMA, Grey Theory, Support Vector Machine and Neural Network. They mainly forecast the situation as a whole, failing to consider the specific situation of the elements of the situation. In order to provide a reliable basis for airport managers to make decisions and defensive measures, this paper uses Hidden Markov Model (HMM) time series method to describe the dependence of airport operation situation at different times, predict the changing trend, and analyze its internal situation. The basic idea is to build a Hidden Markov Model, collect external observable state data and train the Hidden Markov Model to predict the level of airport operation situation. If the number of grades is high, then find out the wrong characteristic sequence and take measures to improve the service capacity of the airport.

2 Airport Operation Situation Analysis

Airport operation situation describes the status of airport operation from macro and overall aspects. Airport operation situation mainly embodies in three aspects: flight guarantee, utilization efficiency of core resources and take-off and landing efficiency.

2.1 Definition and Quantitative Expression of Airport Operation Situation

Situation is a holistic and overall situation concept. The airport operation situation is a macroscopic and overall description of the airport operation situation. Airport operation situation value summarizes and integrates the relevant factors that can reflect the influence of airport operation information into a group of meaningful values. It is a numerical value that can characterize the characteristics of the airport operation situation

$$I(w_i, n_i) = \sum_{i=1}^{n} w_i \cdot n_i \tag{1}$$

where I is the situation value, w_i is the weight of the situation factor i, n_i is the value of the corresponding situation factor i, and n is the number of the situation factor.

2.2 Quantitative Classification of Airport Operation Situation

In terms of operation, flight guarantee, utilization efficiency of core resources and take-off and landing efficiency can more directly reflect the situation of airport operation [15]. Wang Man establishes an evaluation index of flight guarantee from five aspects: average deviation value of arrival time, average deviation value of departure time, average deviation value of guarantee completion time, accuracy rate of target remove chocks time and accuracy rate of calculation of take-off time [16]. Core resource utilization efficiency mainly includes runway utilization efficiency, taxiway utilization efficiency and parking space utilization efficiency. Takeoff and landing efficiency is mainly manifested in two aspects: take-off sorties and landing sorties. In this paper, AHP method is used to determine the weight of situation factor i, and the factors describing airport operation situation are shown in Table 1.

We distributed questionnaires to a number of hub airport staff, among which more than 100 were offline questionnaires and 46 were valid. The questionnaire data are analyzed and sorted out to determine the threshold of the division scope. And the result of the classification of each situation factor and situation value is shown in Table 2. In the table, grades 1, 2 and 3 represent three different situation

Table 1 Quantitative index system of airport operation situation and AHP score

Situation factor	AHP score
Average deviation value of arrival time w1	0.0064
Average deviation value of departure time w2	0.0064
Average deviation value of guarantee completion timew3	0.0297
Accuracy rate of target withdrawal time w4	0.0152
Accuracy rate of calculation of take-off time w5	0.0471
Runway utilization efficiency w6	0.0516
Taxiway utilization w7	0.0516
Parking space utilization efficiency w8	0.1550

Table 2 Grading results of situation values and characteristic sequences

Classification index	Grade		
	Grade 1	Grade 2	Grade 3
Average deviation value of arrival time w_1	(0,19]	(19,26]	(26,1440)
Average deviation value of departure time w_2	(0,20]	(20,31]	(31,1440)
Average deviation value of guarantee completion time w_3	(0,20]	(20,25]	(25,1440)
Accuracy rate of target withdrawal time w_4	(0.85,1)	(0.77,0.85]	(0,0.77]
Accuracy rate of calculation of take-off time w_5	(0.85,1)	(0.77,0.85]	(0,0.77]
Runway utilization efficiency w_6	(0.94,1)	(0.90,0.94]	(0,0.90]
Taxiway utilization w_7	(0.94,1)	(0.90,0.94]	(0,0.90]
Parking space utilization efficiency w_8	(0.94,1)	(0.90,0.94]	(0,0.90]
Take-off sorties w_9	(0,72]	(72,121]	(121,400)
Landing sorties w_{10}	(0,72]	(72,121]	(121,400)
Situation value	(0,49]	(49,81]	(81,316)

levels of airport operation respectively. The smaller the grade, the better the service capability. In order to simplify the calculation and reduce the complexity, the airport operation situation factor values are also selected according to three grades:1, 2, 3. The method of ranking is similar to the method of dividing three types of airport operation situation into intervals. The smaller the deviation, the smaller the grade. The higher the accuracy and utilization rate, the smaller the grade. The more sorties are taken off and landed, the smaller the grade. The results of grading and the description of grading are shown in Tables 2 and 3, respectively.

Table 3 Grade description of airport operation situation

Situation grade	Description
Situation 1	Airport operation is unimpeded and not affected by any link
Situation 2	Airport operation is slightly poor, some links interact with each other, and have little impact on passengers
Situation 3	Airport operation is not smooth, which affects the normal operation of the airport and brings inconvenience to passengers

3 Establishment of Airport Operation Situation Prediction Model

Hidden Markov Model (HMM) is a statistical model based on Bayesian decision theory. It requires strict data structure, but the results of calculation are trustworthy, and the sequence of data formation is highly correlated with time. It can not only visualize the observable objects, but also excavate the hidden research objects, such as the situation grade of airport operation situation.

3.1 HMM Model Design

In this paper, we extract 10 airport operational situation factor sequences, namely, the corresponding average deviation value of arrival time, average deviation value of departure time, average deviation value of guarantee completion time, accuracy rate of target withdrawal time, accuracy rate of calculation of take-off time, runway utilization efficiency, taxiway utilization efficiency, parking space utilization efficiency, take-off sorties and landing sorties. The airport operation situation grade 1, 2 and 3 are characterized. A complete HMM model can be represented by a quintuple. The corresponding parameters are explained as follows:

(1) Because the airport operation situation grades 1, 2 and 3 are not observable, that is, the hidden state in HMM system. The number of hidden states is expressed by parameter $N = 3$, where grade 1, grade 2 and grade 3.
(2) The observed characteristics reflect the output of the system. At any time, the system may have a variety of features observed, recorded as $V = \{v_1, v_2, \ldots, v_M\}$. With the change of time, an observation sequence $O = \{o_1 o_2, \ldots, o_T\}$ can be obtained, where o_t is an object of observation value in time t set V, i.e. $o_t \in \{v_1, v_2, \ldots, v_M\}$. In this paper, the number of observable states is the possible values of 10 airport operational state sequences, so $M = 10$.
(3) A represents the transfer matrix composed of three hidden states of situation grade 1, 2 and 3. $A = (a_{ij})$, where $a_{ij} = P(q_{t+1} = S_j | q_t = S_i)$, $1 \le i, j \le N$, generally speaking, the value of A has little influence on the prediction effect of the model. A specific value can be obtained by running HMM program.

(4) B denotes the probability of the transition from an implicit state to an observable state at a given time, i.e. the confusion matrix, which can also be called the transition probability matrix of the observable state. For example, the group of observational probability distributions of the state is $b_j(k) = P\left(o_t = v_k | q_t = S_j\right), 1 \leq j \leq N, 1 \leq k \leq M$. Note again that the observed probability matrix is composed of 10 parameters, such as the corresponding average deviation value of arrival time, average deviation value of departure time, average deviation value of guarantee completion time, accuracy rate of target withdrawal time, accuracy rate of calculation of take-off time, runway utilization efficiency, taxiway utilization efficiency, parking space utilization efficiency, take-off sorties and landing sorties.

(5) π denotes the initial state probability, $\pi = \{\pi_i\}$,where $\pi_i = P(q_1 = S_i), 1 \leq i \leq N$. It is the initial probability matrix of a certain operational situation grade, which can be obtained from the basic training data. From probability theory, it is easy to find the initial probability matrix.

3.2 Solution Method

Implementing HMM prediction model is divided into two parts: parameter learning and prediction. The basic idea of parameter learning is to use Baum-Welch algorithm (forward-backward algorithm) to calculate $\chi = \left\{O^k\right\}_{k=1}^{K}$, which means K Sequence Samples for maximizing airport operational situation grade to get likelihood HMM model $\hat{\lambda}$, that is, to find out the model $P(\chi|\lambda)$ that maximizes $\hat{\lambda}$. From the given observation sequence, the training set χ of the situation grade sample is constructed. Through training learning, the training set χ is most likely to explain the HMM model $\hat{\lambda}$.

3.2.1 Baum-Welch Algorithm

The learning process of Baum-Welch algorithm is essentially a kind of EM method, and one of the characteristics of EM method is that it needs repeated computation. The steps are as follows:

Step 1: when t = 0, some parameters are assigned a priori value, which needs to satisfy the following three specific conditions:

$$\sum_{i=1}^{N} \pi_i = 1; \sum_{j=1}^{N} a_{ij} = 1, 1 \leq i \leq N; \sum_{k=1}^{N} b_j(k) = 1, 1 \leq j \leq N.$$

Step 2: Expectation-step

We Calculate $\xi_t(i, j)$ and $\gamma_t(i)$ when the sum of observations O is known.

$$\xi_t(i,j) = P\left(q_t = s_i, q_{t+1} = s_j | O, \mu\right) = \frac{\alpha_t(i) * a_{ij} * b_j(O_{t+1}) * \beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i) * a_{ij} * b_j(O_{t+1}) * \beta_{t+1}} \tag{2}$$

$$\gamma_t(i) = \sum_{j=1}^{N}\xi_t(i,j) \tag{3}$$

where

$$\alpha_t(j) = \left[\sum_{i=1}^{N}\alpha_{t-1}(i)a_{ij}\right] \cdot b_{jt} \tag{4}$$

$$\beta_t(j) = \sum_{i=1}^{N} a_{ij} \cdot b_{j_{t+1}} \cdot \beta_{t+1}(j) \tag{5}$$

Step 3: Maximization step

The K sequence samples observed are recorded as $\chi = \left\{O^k\right\}_{k=1}^{K}$. We also assume that they are independent and identically distributed, then the formula is $P(\chi|\lambda) = \prod_{k=1}^{K} P\left(O^k|\lambda\right)$. Finally, the sum of all observable sequences that make up the situation grade is divided by the total number of observable sequences. Repeat the following three calculations until the requirements are met:

$$\hat{\pi}_i = \frac{\sum_{k=1}^{K}\gamma_1^k(i)}{K}; \hat{a}_{ij} = \frac{\sum_{k=1}^{K}\sum_{t=1}^{T_K-1}\xi_t^k(i,j)}{\sum_{k=1}^{K}\sum_{t=1}^{T_K-1}\gamma_t^k(i)}; \hat{b}_{jm} = \frac{\sum_{k=1}^{K}\sum_{t=1}^{T_K-1}\gamma_t^k(j)\cdot\left(O_t^k = v_m\right)}{\sum_{k=1}^{K}\sum_{t=1}^{T_K-1}\gamma_t^k(i)}.$$

3.2.2 Viterbi Algorithm

The estimated value of model $\widehat{\lambda} = (A, B, \pi)$ is obtained by training samples. Given the observation sequence, we can use Viterbi algorithm to solve the prediction problem of hidden Markov model. Viterbi algorithm can solve the maximum probabilistic path by dynamic programming. The model with the largest likelihood logarithm is the corresponding situation grade, and each path corresponds to a most probable state sequence.

Set two variables δ, ψ and use formula (6) to calculate the maximum possibility of a path $(i_1, i_2, \ldots, i_t)$. The assumption is that the time is t and the state is i:

$$\delta_t(i) = \max_{i_1,i_2,\ldots,i_{t-1}} P(i_t = i, i_{t-1}, \ldots, i_1, o_{t+1}, \ldots, o_1|\lambda) \tag{6}$$

According to the above, further deduce t + 1 moment:

$$\begin{aligned}\delta_{t+1}(i) &= \max_{i_1,i_2,\ldots,i_t} P(i_{t+1}=i, i_{t-1},\ldots,i_1, o_{t+1},\ldots,o_1|\lambda) \\ &= \max_{1\le j\le N}\left[\delta_t(t)a_{ji}\right]b_i(o_{t+1}) \end{aligned} \tag{7}$$

Under the same premise and assumption, the path $(i_1, i_2, \ldots, i_{t-1}, i)$ with the greatest possibility can be obtained. And T-1 node formula is as follows:

$$\psi_t(i) = \arg\max_{1\le j\le N}\left[\delta_{t-1}(j)a_{ji}\right] \tag{8}$$

The principle and calculation process of Viterbi algorithm are as follows:

Input: $\overline{\lambda} = (\pi, A, B)$ and $O = (o_1, o_2, \ldots, o_T)$

Output: the most likely sequence is $I^* = \left(i_1^*, i_2^*, \ldots, i_T^*\right)$

Step 1: When t = 1, assign the initial value to the variable: $\delta_1(i) = \pi_i b_i(o_1)$, $\psi_1(i) = 0$

Step 2: When t = t + 1, the loop proceeds: for $t = 2, 3, \ldots, T$

$$\delta_t(i) = \max_{1\le j\le N}\left[\delta_{t-1}(j)a_{ji}\right]b_i(o_t) \tag{9}$$

$$\psi_t(i) = \arg\max_{1\le j\le N}\left[\delta_{t-1}(j)a_{ji}\right] \tag{10}$$

Step 3: End the cycle and get the result:

$$P^* = \max_{1\le i\le N}\delta_T(i) \tag{11}$$

$$i_T^* = \arg\max_{1\le i\le N}[\delta_T(i)] \tag{12}$$

Step 4: Look back for the best path to choose: $t = T-1, T-2, \ldots, 1$, $i_t^* = \psi_{t+1}\left(i_{t+1}^*\right)$. The best sequence of states obtained is $I^* = \left(i_1^*, i_2^*, \ldots, i_T^*\right)$.

4 Example Analysis

The selected hub airport has two terminal buildings T1 and T2, and the airport operation is mainly T2 terminal, with an area of about 500,000 m^2. The runway system of the hub airport is in a relatively advanced state. It is the 4F level that most airports are striving for at present. The length of the runway is nine times as long as the University playground, which is not much different from other hub airports. Although the apron of the hub airport is basically provided for passenger aircraft, there are several vendors which can basically meet the needs of passenger

transport. In addition, the junction between the hub airport and the ground traffic is relatively good. There are other infrastructures, such as subway and bus stations, for traffic users to reach their destinations. The corresponding supporting facilities are relatively complete and can accommodate all types of aircraft. Therefore, the hub airport as a case study is representative of the hub airport.

4.1 Data Collation and Analysis

According to the case of hub airport, we analyses the status of airport operation, and evaluates the capability of airport operation and service. The basic data of Xinzheng Airport from April 1 to April 30, 2016 were collected, including the actual arrival time of flights, planned arrival time, actual departure time, planned departure time, target departure time, target block time, target withdrawal time, runway capacity, taxiway capacity and parking capacity. After calculating and sorting out, 30 groups of training data were obtained (including inbound time deviation value, departure time deviation value, guarantee completion time deviation value, target withdrawal time accuracy rate, planned take-off time accuracy rate, runway utilization rate, taxiway utilization rate, parking space utilization rate, take-off and landing sorties). Using 10 sets of data from May 1 to May 10 as test data (i.e. forecast data). A total of 40 groups of situation values and rank sequence values of training data and test data are shown in Fig. 1 and Table 4, respectively.

Figure 1 shows that the fluctuation of airport operation situation value in each period is not significant. The fluctuation of airport operation situation value varies from 120 to 160. The lowest value appears on May 3 and the highest value appears on April 30. Because the principle of grading is that The principle of grading is that the

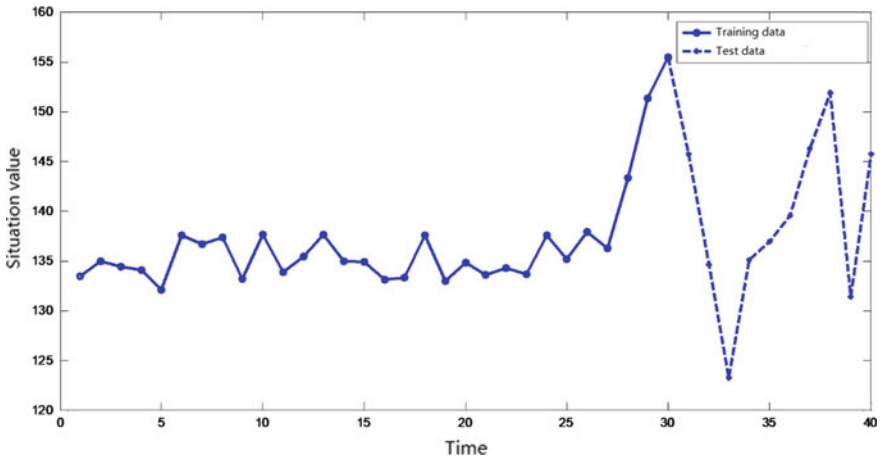


Fig. 1 Situation values of training data and test data

Table 4 Hierarchical sequence values of basic data

Date	Grade										
	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	Situation value
4.1	1	1	2	1	1	3	3	3	2	2	2
4.2	3	3	3	1	1	3	3	3	3	3	3
4.3	2	2	1	1	1	3	3	3	2	2	2
4.4	2	1	2	1	1	3	3	3	2	2	2
4.5	1	1	1	1	1	3	3	3	2	2	2
4.6	3	3	1	3	2	2	2	2	3	3	3
4.7	2	2	1	3	3	2	2	2	2	2	2
4.8	2	1	2	3	3	1	1	1	2	2	2
4.9	3	3	3	2	3	3	3	3	2	2	2
4.10	1	2	1	1	1	3	3	3	3	3	3
4.11	2	3	1	2	2	2	2	2	2	2	2
4.12	2	2	3	1	1	2	2	2	3	3	3
4.13	1	2	2	3	3	1	1	1	2	2	2
4.14	1	1	2	3	3	1	1	1	2	2	2
4.15	1	1	3	3	2	1	1	1	2	2	2
4.16	2	1	3	1	1	1	1	1	2	2	2
4.17	1	1	2	1	1	1	1	1	3	3	3
4.18	2	2	3	3	3	1	1	1	3	3	3
4.19	1	2	1	2	1	1	1	1	2	2	2
4.20	2	3	2	1	2	3	3	3	3	3	3

(continued)

Table 4 (continued)

Date	Grade										
	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	Situation value
4.21	3	3	1	1	2	1	1	1	2	3	2
4.22	3	2	1	3	2	2	2	2	3	3	3
4.23	2	2	1	3	3	2	2	2	2	2	2
4.24	1	2	1	3	3	1	1	1	3	3	3
4.25	1	3	3	2	3	3	3	3	2	2	2
4.26	1	2	1	1	1	3	3	3	2	2	2
4.27	3	2	1	2	2	2	2	2	2	2	2
4.28	2	2	2	1	1	2	2	2	2	2	2
4.29	2	2	3	3	3	1	1	1	1	1	1
4.30	1	1	2	3	3	1	1	1	1	1	1
5.1	3	3	3	3	2	1	1	1	2	1	1
5.2	2	2	3	1	1	3	3	3	3	1	2
5.3	2	2	3	1	1	3	3	3	3	3	3
5.4	3	3	3	1	1	3	3	3	3	1	2
5.5	2	2	3	3	2	2	2	2	3	1	2
5.6	3	3	3	1	1	1	1	1	3	1	1
5.7	3	3	3	3	3	1	1	1	2	1	1
5.8	3	3	3	3	3	1	1	1	1	1	1
5.9	3	3	3	2	2	2	2	2	1	3	3
5.10	3	3	3	3	3	1	1	1	2	1	1

greater the situation value, the better the airport service. The main reason for finding the lower situation value on May 3 is that there are fewer aircraft take-off sorties. Considering the actual situation, May 1, 2016 is a holiday. The period ended on May 2, resulting in fewer sorties than in the previous two days or even in peace time. Similarly, the highest value appeared on April 30, which shows that the utilization rate of airport is higher, the number of people traveling on the day of holiday is larger, and the number of sorties taking off and landing is larger, so the situation value is larger. It can be seen that the most important factor affecting the airport operation situation is the take-off and landing sorties, which reflects the larger weight of the former two factors. From the grade, most of them are in situation grade 2 and grade 3, which indicates that the efficiency of airport operation needs to be improved.

4.2 Verification of HMM Prediction Model

According to the basic data of airport operation situation value, the main program of target is compiled and the HMM model is trained by MATLAB. K sample sequences $\chi = \left\{O^k\right\}_{k=1}^{K}$ are used to train the hidden Markov model until the result converges, and the likelihood model $\widehat{\lambda}$ is obtained. Through training, the following three groups of likelihood model parameters with situation grades of 1, 2 and 3 are obtained, and their learning curves are shown in Fig. 2.

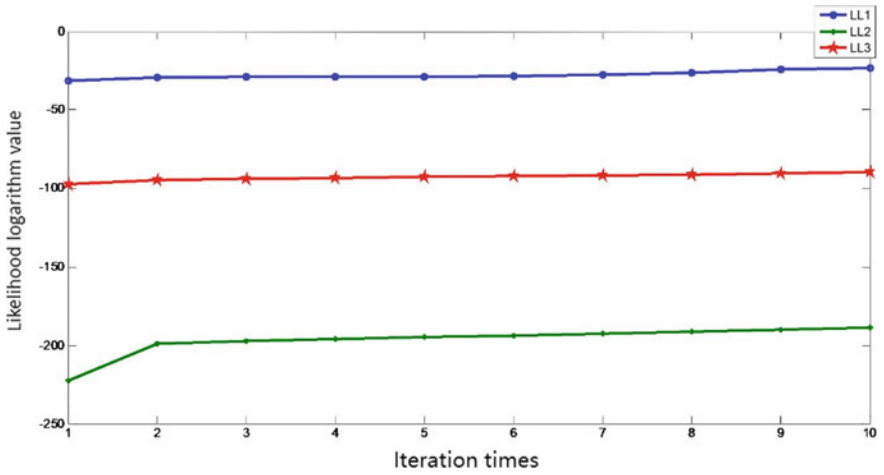


Fig. 2 The learning curve of HMM with the number of iterations

$$\hat{\lambda}_1 = (\pi_1, A_1, B_1)$$

$$\pi_1 = (0.5572, 0.0789, 0.3640), A_1 = \begin{pmatrix} 0.2296 & 0.3503 & 0.4201 \\ 0.4641 & 0.2469 & 0.2890 \\ 0.2058 & 0.3525 & 0.4416 \end{pmatrix}, B_1 = \begin{pmatrix} 0.4375 & 0.4928 & 0.0697 \\ 0.6041 & 0.0604 & 0.3355 \\ 0.1369 & 0.5100 & 0.3530 \end{pmatrix}$$

$$\hat{\lambda}_2 = (\pi_2, A_2, B_2)$$

$$\pi_2 = (0.7366, 0.1163, 0.1472), A_2 = \begin{pmatrix} 0.4559 & 0.2461 & 0.2980 \\ 0.4059 & 0.3477 & 0.2464 \\ 0.1511 & 0.3538 & 0.4951 \end{pmatrix}, B_2 = \begin{pmatrix} 0.2233 & 0.3658 & 0.4109 \\ 0.4467 & 0.0061 & 0.5472 \\ 0.4290 & 0.0839 & 0.4872 \end{pmatrix}$$

$$\hat{\lambda}_3 = (\pi_3, A_3, B_3)$$

$$\pi_3 = (0.0807, 0.6269, 0.2925), A_3 = \begin{pmatrix} 0.2231 & 0.542 & 0.2328 \\ 0.2018 & 0.7196 & 0.0786 \\ 0.4336 & 0.3006 & 0.2657 \end{pmatrix}, B_2 = \begin{pmatrix} 0.3498 & 0.4275 & 0.2227 \\ 0.4080 & 0.3754 & 0.2166 \\ 0.5623 & 0.0975 & 0.3402 \end{pmatrix}$$

Using MATLAB software, 10 groups of test sample data from May 1 to May 10 (including five groups of situation grade 1, three groups of situation grade 2, two groups of situation grade 3) are substituted into the three HMM models trained above. And the corresponding logarithmic likelihood values of these 10 groups of test samples are calculated. The maximum likelihood value is the model with the highest matching degree. The predicted results of airport operation situation grade and the most possible state sequence corresponding to the next time are shown in Tables 5 and 6 respectively. The gray labeling is the corresponding situation grade.

On this basis, the Accuracy, Mean Absolute Percent Error (MAPE), Root Mean Square Error (RMSE) and Relative Error (RE) of the model are tested. Compared with the Autoregressive Moving Average model ARMA and Grey Markov model GM, which are traditional situation prediction methods, the HMM model can better

Table 5 Airport operation situation prediction results

Test samples	Situation model		
	Situation grade 1	Situation grade 2	Situation grade 3
Grade 1 Sample 1	−11.3553	−11.7723	−16.0799
Grade 1 Sample 2	−8.4170	−11.8053	−9.1466
Grade 1 Sample 3	−10.4883	−12.3628	−10.8605
Grade 1 Sample 4	−8.5825	−11.8038	−9.4475
Grade 1 Sample 5	−9.6441	−11.3983	−11.2392
Grade 2 Sample 1	−11.2593	−11.1399	−14.3656
Grade 2 Sample 2	−10.7502	−6.8765	−12.5598
Grade 2 Sample 3	−14.0003	−11.0104	−12.7112
Grade 3 Sample 1	−12.9649	−12.3707	−10.1651
Grade 3 Sample 2	−14.0942	−29.6147	−11.6856

Table 6 Most possible state sequences for airport operational situation grade forecast in the next time

Date	Situation factor									
	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10
5.1	1	1	1	1	2	3	2	3	2	3
5.2	2	3	1	1	3	3	3	3	1	2
5.3	2	3	1	1	3	3	3	3	3	2
5.4	1	1	2	3	3	1	2	1	2	3
5.5	2	2	3	3	2	2	2	2	3	1
5.6	3	1	2	1	2	1	2	1	2	1
5.7	3	3	3	3	3	2	3	2	3	2
5.7	1	3	1	3	1	3	2	1	3	2
5.9	2	2	2	1	3	1	3	1	2	2
5.10	1	1	1	1	2	3	3	3	2	3

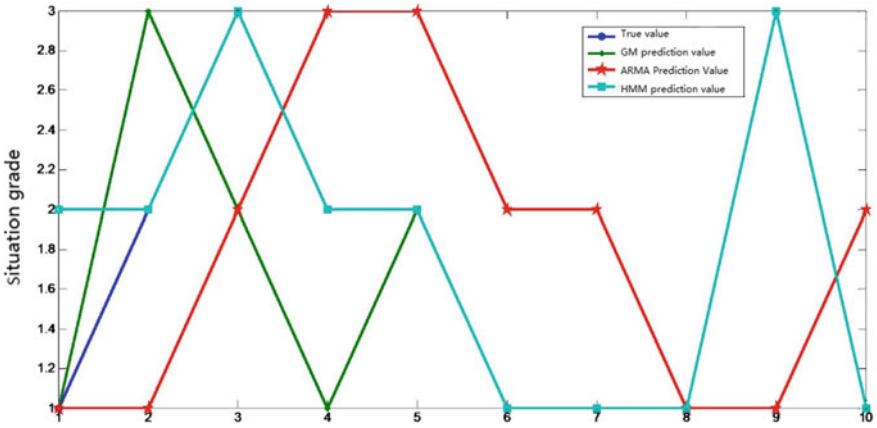


Fig. 3 Predicted values of three different methods of rank series

Table 7 Accuracy and error comparison results of different models

Contrastive terms	Model		
	ARMA	GM	HMM
Accuracy (%)	30	70	90
MAPE (%)	5.981	3.312	3.189
RMSE	9.7850	7.6081	6.8828
RE	0.0005	0.003	0.0003

describe the operation situation of the airport, which further verifies the correctness of the HMM model. The predicted values are shown in Fig. 3, and the Accuracy and Error Results are shown in Table 7.

4.3 Application of Prediction Model

According to the forecast result of situation grade in Fig. 3, we can easily find that the operation situation grade of the airport is 3 on May 3 and 9. Comparing with the corresponding rank sequence in Table 4, we can find out the reasons that affect their high rank. As shown in Table 8, we can take some measures to prevent them in advance. For example, on May 3, attention should be paid to air control to reduce departure value of arrival time, shorten guarantee completion time, dynamically deploy to improve the utilization of runway and parking space. And landing sorties are related to many factors, which can not be avoided simply through the details of airport staff. Similarly, the defensive measures of May 9 can be obtained. This

Table 8 Main reasons for high grades on 3 May and 9 May

Date	Main reasons
May 3	Large w_1, Large w_3, Low w_6, Low w_8, Few w_{10}
May 9	Low w_6, Low w_7, Low w_8, Few w_{10}

prediction result can be used as a theoretical basis for airport staff to pay attention to a certain link. Therefore, it has certain practicability.

5 Conclusion

(1) Using Hidden Markov Model to predict the airport operation situation grade can effectively grasp the quality of airport operation service ability to make coping decisions;
(2) Comparing with the traditional situation prediction methods such as ARMA and GM, it is found that the model adopted in this paper has higher prediction accuracy and smaller error, and can provide improvement measures to improve the service capacity of the airport and better fit the operation of the airport.
(3) The division of the whole operation process of the airport is not meticulous enough. The selected characteristic states are relatively macroscopical. Follow-up research can carry out micro-research on Airport operation, and combine macro-and micro-research to provide theoretical support for airport operation situation prediction.

Acknowledgements This research has been supported under the National Natural Science Foundation of China (Grant No. U1533203), Sichuan Science and Technology Project (Grant No. 2019YFG0050), Sichuan Provincial-College Cooperation Science and Technology Project (Grant No. 2019YFSY0024).

References

1. Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Hum Factors 37(1):32–64
2. Burns CM, Skraaning G, Jamieson GA et al (2008) Evaluation of ecological interface design for nuclear process control: situation awareness effects. Hum Factors J Hum Factors Ergon Soc 50(4):663–679
3. Panteli M, Kirschen DS (2015) Situation awareness in power systems: theory, challenges and applications. Electr Power Syst Res 122:140–151
4. Nonose K, Corver S, Majumdar A et al (2014) A behavioral observation method to assess team situation awareness of air traffic control teams. In: Proceedings of the 26th Australian

computer-human interaction conference on designing futures: the future of design. ACM, pp 456–459
5. Blasch E, Wang Z, Shen D et al (2014) Enhanced air operations for ground situational awareness. In: 2014 IEEE/AIAA 33rd digital avionics systems conference (DASC). IEEE, pp 3D2-1–3D2-13
6. Xie L, Wang Y, Yu J (2013) Network security situation awareness based on neural networks. J Tsinghua Univ (Science and Technology) 23(12):1750–1760
7. Xu X, Ren J, Li N (2014) Identification of terminal area traffic situation based on FCM. Aeronaut Comput Tech 44(1):1–4
8. Li N, Ren J, Xu X (2014) Identification of terminal area traffic situation. Sci Technol Eng 14(11):256–261
9. Wei R (2015) Study on the methods of multi-scale prediction for expressway. Jilin University
10. Wang G (2016) Researches on flight status estimation and display based on GIS. Civil Aviation University of China
11. Kun Qian (2016) Design of traffic operation situation analysis system based on intelligent city intelligent transportation system. China Manage Informationization 19(21):208–210
12. Yuan L (2017) Research on dynamic traffic characteristics and operation situation of terminal area. Nanjing University of Aeronautics and Astronautics
13. Feng C, Jing X, Li Q, Yao P (2017) Theoretical research of decision-making point in air combat based on hidden Markov model. J Beijing Univ Aeronaut Astronaut 43(3):615–626
14. Xu C, Jiang Y, Cai M, Chen L (2018) Joint Scheduling of both taxiway and gate based on bi-level programming. J Beijing Univ Aeronaut Astronaut
15. Yang W, Du Z, Zhou Y (2017) Overview of airport operation situation awareness system and key technologies. In: 2017 Papers of the world transport congress
16. Wang M (2016) Research on association rule discovery and model of warning evaluation about flight cooperative security. Civil Aviation University of China

Parking Demand Forecasting Model for Urban Complex Based on Shared Parking: A Case Study of Harbin City

Xian-cai Jiang and Longyang Zhang

Abstract The paper takes urban complex as research object, aiming at the problem that the current parking facilities allocation neglects the difference of parking demand features of various buildings, which worsens the contradiction between supply and demand of parking. Analysis shows that the peak parking hour of various buildings in the urban complex is complementary, and the supply and demand are seriously imbalanced based on the parking survey data in Harbin. The main influence factors of urban complex parking demand are analyzed, and the revision coefficient of parking generation rate model of urban complex under the influence of a single factor is constructed combining with the actual parking demand. Based on the idea of shared parking, a parking demand forecasting model of urban complex under the comprehensive action of multiple factors is established by using regression analysis method, and the Yuguang-Intel Industrial Park in Harbin is taken as an example to verify the validity of the model. The results show that the predicting value of parking demand by the model is closer to the actual parking demand, which can effectively avoid the imbalance between supply and demand, and improve the utilization efficiency of parking facilities.

Keywords Traffic engineering · Parking demand forecasting · Regression analysis · Urban complex · Shared parking

1 Introduction

The idea of shared parking was first applied to practice in Portland city, USA [1]. Later, the shared parking model was proposed at the Tamper City Conference, which laid the foundation for the allocation of shared parking resources in the future [2]. In view of the phenomenon of monopoly in parking market, Mary put forward a method to modify parking demand with non-monopoly coefficient [3]. Jiang et al.

X. Jiang (✉) · L. Zhang
School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China
e-mail: jxc023@126.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_76

[4] established a new parking demand model based on parking generation rate model, taking into account the distance to the destination after parking, parking search time, parking time and parking cost. Xie et al. proposed a bi-level programming model for university shared parking. The upper level was the four criteria to implement shared parking, and the lower level was the improved SEM-Logit model [5]. Ran et al. [6] combined with parking demand management strategy and shared parking implementation process in China, constructed a shared parking framework. Xue et al. [7] established a parking demand prediction model adapted to new cities by analogy method, considering the revision of parking generation rate model in planning year by peak parameter and utility reduction function of shared parking. Li [8] put forward the corresponding countermeasures for the feasibility, conditions, management mode and steps of shared parking. Su et al. [9] revised the parking demand forecasting model of Chinese cities by using non-monopoly coefficient. Xu and Chen [10] proposed a bi-level programming model to allocate dynamic traffic demand for shared parking. Duan et al. [11] established a model to evaluate the capacity of community parking to share with the outside world. The study showed that residential areas could provide 55% of the shared parking for adjacent buildings.

At present, the parking requirements of different cities in China takes into account the differences of peak parking demand of different types of buildings, but they are all based on a single category of buildings. For urban complexes with various types of buildings, the peak parking demand of different types of buildings is only accumulated when parking demand is calculated. This treatment not only enlarges the contradiction between supply and demand of parking space, but also leads to low efficiency of parking space turnover. Based on this, the parking demand forecasting method of urban complex is studied.

2 Parking Demand Features of Urban Complex

2.1 Definition

Urban complex refers to urban construction projects composed of two or more land utilities, and different types of buildings. Different buildings share space and are complementary in value [12].

2.2 Analysis of Parking Demand Features

(1) Wanda Plaza in Harbin West Railway Station

As a peripheral urban complex of residence, commerce, catering and other functions, it is equipped with 1386 parking spaces. The parking demands in different categories of buildings at various time are shown in Fig. 1.

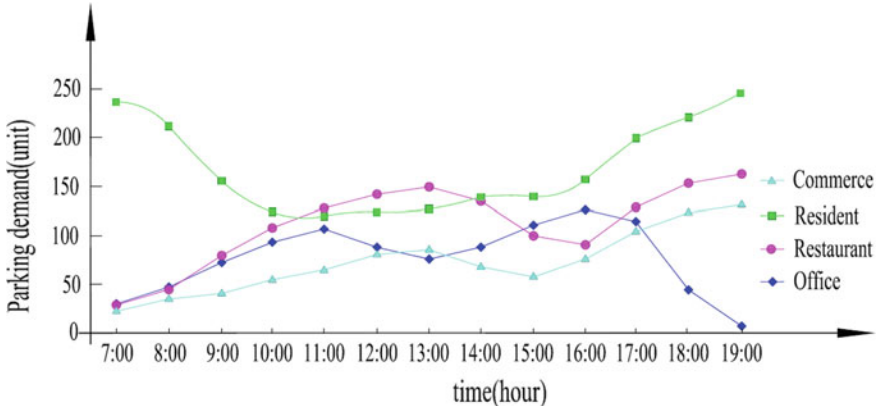


Fig. 1 Parking demand changes in different types of buildings

As can be seen from Fig. 1, the peak parking hours of different types of buildings are different and complementary. The ones of residential buildings are at night, while those hours of catering, office and commercial buildings are complementary with residential buildings.

(2) Qiulin business district

It is a complex in the down town with commercial, residential, office functions enclosed with Bank Street, Ashi River Street, Garden Street and Hongjun Street. Qiulin district should be collocated with 3408 new parking spaces to meet current demands, in accordance with the "Interim Provisions for the Planning and Construction of Parking Lots Collocating Construction Projects in Harbin and Public Parking Lots (2009)". A survey on parking shows that only 2895 parking lots can meet the demand in peak hour, as seen in Fig. 2.

Therefore, the actual parking demands of urban complex are inconsistence with collocating construction requirements calculated from the parking generation rate model, rooting in that the collocating construction of parking lots is calculated by the accumulation of peak parking demands in various buildings, ignoring the difference of peak parking hours in different buildings.

3 Influence Factors on Parking Demand in Urban Complex

(1) Location

Different from the single type of building, the influence of location conditions on parking demand of the urban complex is not strictly in accordance with the increase of the distance from CBD, and is closely related to the travel purpose which the urban

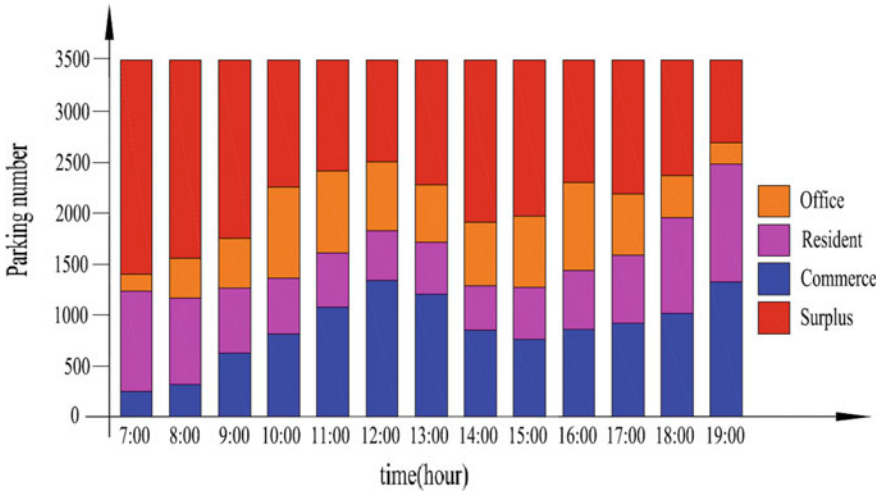


Fig. 2 The gap between parking demand and parking supply for urban complex

complex can bring about. Therefore, the location conditions have great influence on the urban complex parking demand.

(2) Building scale

The parking demand of a single kind of building increases with the building scale, while the traffic attraction of the urban complex is stronger than that of the single class building, which results in the increase of parking demand.

(3) Mixing degree

Urban complex consists of different types of buildings. The impact on parking demand will be more prominent with the more building categories and the more traffic travel purpose which can be achieved. The number of different types of buildings in urban complex is indicated by the mixing degree.

(4) Mixing ratio

When the proportion of different types of buildings is not the same as for the same scale of the urban complex, which is said that different types of buildings have different construction scale. It has a direct impact on the change in parking demand.

(5) City scale

The corresponding parking requirements will be greater with the larger city size and the greater agglomeration effect of the traffic. But this paper only studies the parking demand of a particular city, the influence of the city size can be excluded.

(6) Other factors

Parking capacity and policy will also have an impact on the urban complex parking demand. Similarly, this paper only studies the parking demand of a particular city, which can be ignored.

4 Parking Demand Forecasting Model Based on Shared Parking

4.1 Definition of Shared Parking

Shared parking refers to that different types of buildings can share parking spaces each other due to the different peak parking hours in a certain area [13].

4.2 Quantification of Influence Factors

The correction coefficient for single factor is the ratio of actual parking demand under this influence factor, including illegal parking, and the parking number calculated according to the construction index.

(1) Location (λ_1)

In this paper, the city central area is within the second ring in Harbin, the transition zone is outside the second ring and within the third ring, and the suburbs is outside the third ring. Due to the change of parking demand with time, the total parking demand must be multiplied by a correction coefficient which referred to 1.1 [14] to compensate for the impact of the unsynchronized investigation periods. The relationship between the actual parking demand of different urban complex and the number of parking spaces calculated according to the construction index is shown in Table 1.

The maximum correction coefficient of each location is taken as the correction coefficient of the parking generation rate model of the location condition, and the results from Table 1 are shown as follows:

$$\lambda_1 = \begin{cases} 1.20 \text{ Central area} \\ 1.05 \text{ Transition zone} \\ 0.85 \text{ Suburb} \end{cases} \tag{1}$$

(2) Construction scale (λ_2)

The relationship between the actual parking demand and the number of parking spaces calculated by the construction index of urban complex in Harbin is shown in Table 2 and Fig. 3.

Table 1 Correction of location conditions on the parking generation rate model

Name of urban complex	Location	Total parking demand (unit)	Parking spaces of construction (unit)	Correction coefficient
International Convention Centre	Central area	2597	2220	1.17
Modern Culture Art Garden	Central area	1908	1767	1.08
Taixin Canon	Central area	2508	2161	1.16
Hongze Central Park	Central area	2255	1879	1.20
East Wanxiangshang	Central area	2417	2031	1.19
Boyue Star Town	Central area	2039	1789	1.14
Xiangfang Shanty Town	Central area	2277	2070	1.10
Railway Goods Yard in Haxi	Transition zone	2104	2238	0.94
Yuanda in Qunli	Transition zone	1711	1661	1.03
Tianyue International	Transition zone	2095	1995	1.05
East Zhixin Yuelan	Transition zone	2383	2314	1.03
Golden Ruilin Town	Transition zone	2052	2094	0.98
Baorong Carol International	Transition zone	1946	2093	0.93
Huarun Triumphal Arch	Transition zone	1714	1785	0.96
Yulong Bay	Suburb	1186	1463	0.81
Shimao Dublin	Suburb	1180	1476	0.80
Zhongguan International	Suburb	1455	1712	0.85
Golden Estuary	Suburb	1658	2099	0.79
Meiyu River Island	Suburb	1143	1412	0.81
Jintai Green Lakeside	Suburb	1153	1498	0.77

Table 2 Correction of building scale on the parking generation rate model

Name of urban complex	Construction area (m^2)	Total parking demand (unit)	Parking spaces of construction (unit)	Correction coefficient
Bainian Lijing	180,000	936	1246	0.75
Hengda Dijing	210,000	1446	1854	0.78
Jin'an Sunflower	270,000	1588	2010	0.79
Walking in Paris	390,000	1892	2307	0.82
Rose Bay in Qunli	420,000	2094	2464	0.85
Royal Gardens	530,000	3293	3700	0.89
Guanjiang International in Qunli	620,000	2839	3086	0.92
Hongrun Lakeville	800,000	4030	4198	0.96
Jiangwan Town	880,000	4336	4380	0.99
Binjiang Manufacturers	1,100,000	5418	5160	1.05
Star Square	1,480,000	7633	7068	1.08
Yintai Town	2,320,000	11,506	10,273	1.12
Lesong Business Area	3,240,000	19,761	16,890	1.17
Tongxiang Business Area	4,340,000	25,061	21,060	1.19
Haxi Business Area	5,120,000	34,046	27,680	1.23

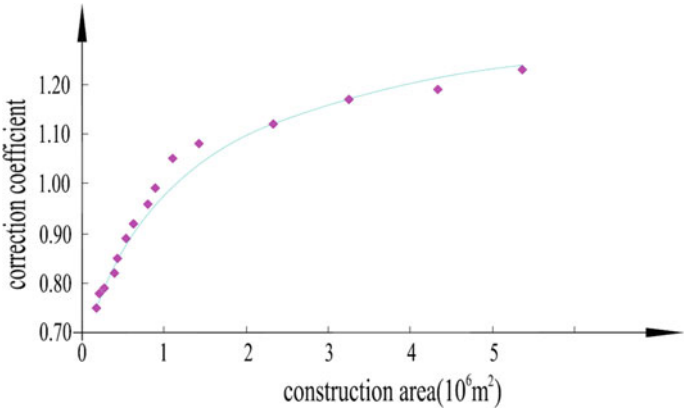


Fig. 3 Correction relationship of building scale on the parking generation rate model

$$\lambda_2 = 0.1464 \ln s + 0.9955,\ R^2 = 0.9835 \tag{2}$$

where s is referred to the construction area of urban complex.

As can be seen from Fig. 3, the correction coefficient will be greater with the larger size of the urban complex construction and the greater aggregation effect of traffic. The correction coefficient increases rapidly in the beginning. When the building scale reaches a certain extent, the growth trend gradually slows down.

(3) Mixing degree (λ_3)

The mixing degree 2 refers to the urban complex composed of two kinds of buildings, such as residential and commercial buildings, and the mixing degree 3 refers to the urban complex composed of three types of buildings. Each kind of mixing degree selects 3 urban complexes. After the correction coefficient and its' average value are calculated, the correction coefficient will be determined through the curve fitting, as seen in Table 3 and Fig. 4.

$$\lambda_3 = 0.0045h^2 + 0.0484h + 0.7221,\ R^2 = 0.9766 \tag{3}$$

where h is the number of urban complex types.

The correction coefficient of parking generation rate model increases with the increase of the urban complex mixing degree from Fig. 4.

(4) Mixing ratio (λ_4)

Even though some urban complex has the same mixing degree, their mixed species vary widely. For example, mixing degree 2 can be residence and commerce, or can also be commerce and office. The mixing ratio, as the ratio of construction area, is complex for each kind of mixing degree. The paper only studies the common mixing degree of 2 (commercial and residential) and 3 (commercial, office and residential).

The results of the urban complex parking demand investigation with different mixed proportions in Harbin are shown in Table 4 and Fig. 5.

As can be seen from Fig. 5, when the mixing ratio is 3:7, the correction coefficient is the smallest, and the effect of the shared parking is the best. With the increasing or decreasing in the proportion of commercial buildings, the condition of shared parking will be worse, which occurs free parking condition.

The parking demand survey results of 13 urban complex mainly residential are shown in Fig. 6.

It can be seen from Fig. 6 that when a kind of building is dominant, the correction coefficient increases with a staggered oscillation trend with the increase of the building. There are many kinds of mixed proportions of the three types of buildings, whose effect on the parking of the construction index is more complex than that of the two types of buildings, and the evolution trend is difficult to determine.

Table 3 Correction of mixing degree on the parking generation rate model

Name of urban complex	Mixing degree	Total parking demand (unit)	Parking spaces of construction (unit)	Correction coefficient
Sanhuan Information Center	2	1410	1678	0.84
5th District of Paris	2	1622	1886	0.86
Xidu Impression	2	1318	1588	0.83
Jinshui Shangdu	3	1830	2080	0.88
Zhongjiao Chansons	3	1706	1854	0.92
Yuanchuang Yuefu	3	1962	2180	0.90
Hengxiang Space	4	2229	2346	0.95
Rome Park	4	2295	2468	0.93
Xianglin Courtyard	4	2158	2248	0.96
Oriental Plaza	5	2754	2504	1.10
Shangdong Splendid Town	5	2534	2586	0.98
Qikaili Garden	5	2594	2674	0.97
Rose Bay in Qunli	6	3293	2768	1.19
Minjiang International	6	3117	2576	1.21
Splendid City	6	3136	2680	1.17
Napa Ying County	7	3259	2546	1.28
Haxi Washington	7	3160	2508	1.26
Yiyuan King Street	7	3060	2468	1.24

4.3 Parking Demand Forecasting Model

The influence of mixing ratio is not considered in view of the complexity and irregularity of mixing ratio on the correction of parking generation rate model.

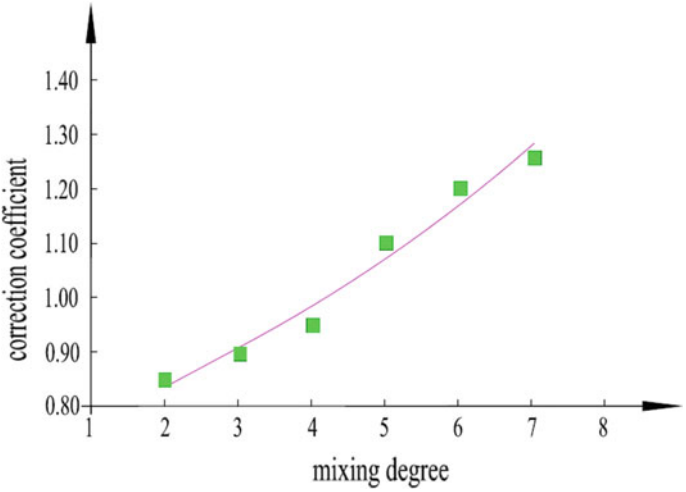


Fig. 4 Correction relationship of mixing degree on the parking generation rate model

Table 4 Correction of mixing ratio on the parking generation rate model (two types of buildings)

Name of urban complex	Mixing ratio	Total parking demand (unit)	Parking spaces of construction (unit)	Correction coefficient
City Star	1:9	1493	1678	0.89
Shengshixiang Bay	2:8	1490	1886	0.79
Mediterranean Sun	3:7	1239	1588	0.78
Europe and Asia world Sunshine	4:6	1664	2080	0.80
View Country	5:5	1539	1854	0.83
Shenghengji Living Mall	6:4	1962	2180	0.86
Huarun Ode to Joy	7:3	1875	2346	0.89
Kaili Square	8:2	2295	2468	0.93
Vienna Park	9:1	2158	2248	0.96

(1) Total correction coefficient

The location, building scale, correction coefficient of mixing degree and the total correction coefficient of 15 urban complexes in Harbin are calculated respectively. The results are shown in Table 5.

The relationship between the total correction coefficient and the single influence factor correction coefficient is gotten as follows by regression analysis from Table 5.

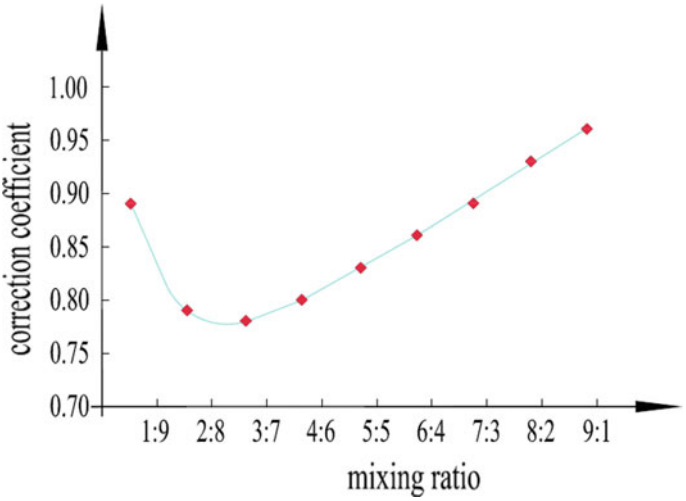


Fig. 5 Correction relationship of mixing ratio on the parking generation rate model

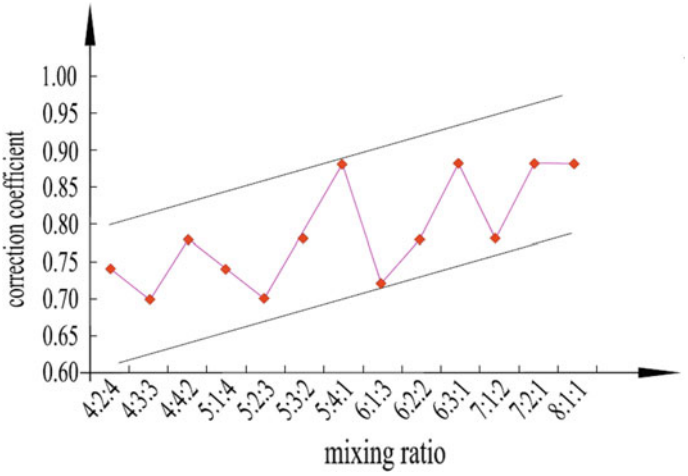


Fig. 6 Correction relationship of mixing ratio on the parking generation rate model

$$\lambda = 0.53\lambda_1 + 0.27\lambda_2 + 0.20\lambda_3,\ R^2 = 0.987 \tag{4}$$

(2) Forecasting model

In summary, the parking demand forecasting model under the multiple influence factors of urban complex based on the parking generation rate model is as follows:

Table 5 The total correction coefficient of parking generation rate model of urban complex

Name of urban complex	Location condition λ_1	Construction scale λ_2	Mixing degree λ_3	Total correction coefficient λ
Fuda Blue Mountain	1.20	0.75	0.86	1.01
Huahong International Center	1.20	0.78	0.88	1.02
Huafeng Haicheng Bay	1.20	0.92	0.96	1.08
Xiangfang Wanda City	1.20	0.98	1.05	1.11
Gelan Yuntian	1.05	0.82	0.83	0.94
Huixiong Time	1.05	0.85	1.10	1.01
Exhibition Town	1.05	0.89	0.95	0.97
Songjiang New City	1.05	0.96	0.90	0.99
Yongtai Town	1.05	1.08	0.86	1.02
Star East Town	0.85	0.75	0.88	0.83
CBD Rui Town	0.85	0.79	0.92	0.85
Wanda Wenlv Town	0.85	0.92	0.95	0.89
City Star	0.85	1.05	0.97	0.93
Huanan Town	0.85	1.23	1.19	1.02

$$D = \lambda \max\left[\sum_{i=1}^{n} R_i^t \cdot S_i\right] \tag{5}$$

where D is the number of parking spaces to be built (unit), R_i^t is the number of parking spaces per unit area of the type i building (unit/100 m^2), S_i is the construction area of the type i building (100 m^2).

5 Case Analysis

5.1 Overview

Hanan Metro Yuguang-Intel Industrial Park is located outside the Third Ring in Harbin, which is composed of residential and office. Its total construction area is 290,302.86 m^2, including the residential area of 145,151.43 m^2, and the office area

of 111,487.33 m^2. The rest are parking lots with 1859 parking spaces and supporting facilities, and the actual maximum parking demand is 1462 vehicles.

5.2 Analysis of Forecasting Results

(a) Calculation of correction coefficient

The project is located outside the 3rd Ring in Harbin, the location correction coefficient is 0.85 calculated by Formula (1). The construction scale correction coefficient of the project is 0.84 calculated by Formula (2). The mixing degree correction coefficient for this project is 0.81 calculated by Formula (3). The total correction coefficient is 0.84 determined by Formula (4).

(b) Analysis of forecasting results

Yuguang-Intel Industrial Park just needs to build 1562 parking spaces to meet the demand under the sharing mode according to the "Regulations" and all kinds of buildings construction scale by Formula (5), while the actual parking demand is 1462 vehicles. The correction coefficient is 1.1, considering the influence of unsynchronized investigation time, and the maximum parking demand calculated is 1608, which decreases by 16% than the parking spaces (1859) determined by a single kind of building and its' construction index. The result indicates that the parking demand forecasting model in this paper is feasible.

6 Conclusion

(1) The analysis shows that the peak parking hours of various buildings in urban complex are complementary. Besides, the contradiction between supply and demand of parking facilities increases and the utilization efficiency of facilities is low.
(2) According to the actual parking demand, the paper constructs the parking generation rate model correction coefficient of urban complex under the single influence factor. And the parking demand forecasting model of urban complex is established under the multiple factors by regression analysis method based on the idea of shared parking, and the case analysis verifies the validity of the model.

Acknowledgements This research is supported by the Harbin Special Fund Program in Innovation Talents of Science and Technology (2016RAQXJ079).

References

1. Stin HS, Resha J (1997) Shared parking handbook. Stein Engineering, Oregon, pp 18–19
2. The City Council of the City of Tempe (2006) Zoning and development code of the city of tempe, Arizona. Planning Department of Tempe, Arizona, pp 20–21
3. Smith MS (2005) Shared parking. Urban Land Institute, Washington, DC, pp 50–59
4. Jiang Y, Peng B, Dai L et al (2011) Parking demand forecasting of urban comprehensive development blocks involving shared parking and location conditions. In: Proceedings of the 3rd international conference on transportation engineering. American Society of Civil Engineers, Chengdu, pp 829–834
5. Xie K, Chen J, Zheng JH (2014) Research on parking demand model of colleges and universities based on shared parking. In: Applied mechanics and materials, vol 505. Trans Tech Publications Ltd., Kunming, pp 415–421
6. Ran JY, Guo XC, Chen YM (2009) A framework for CBD parking demand forecasting. Urban Transp China 7(6):59–65
7. Xue XJ, Ou XQ, Yan KF (2010) Parking demand forecasting for space sharing facility in new urban area. Urban Transp China 8(5):52–56
8. LI F (2012) The strategies for the mode of shared parking resource in settlements. Master's thesis of Dalian University of Technology, Dalian, pp 28–40
9. Su J, Guan ZH, Qin HM (2013) Shared parking facility for mixed parking demand generated by different types of land use. Urban Transp China 11(3):42–46
10. Xu XD, Chen J (2014) The bi-level programming model based on dynamic shared parking policy. In: A new type of urbanization and transportation development—the annual meeting of Chinese urban transportation planning in 2013 and the 27th academic symposium proceedings. Academic Committee of Urban Transportation Planning of Urban Planning Society of China, Beijing, pp 1–11
11. Duan MZ, Yang ZS, Zhang L et al (2015) Residential parking spaces shared capability assessment model. J Transp Syst Eng Inf Technol 15(4):106–112
12. Long GX (2011) The development research and practice of large urban complex. Southeast University Press, Nanjing, pp 65–78
13. Bai Y, Xue K, Yang X (2004) Forecasting method of parking-demand based on capacity-of-network. J Traffic Transp Eng 4(4):49–52
14. Li LB, Wang M, Dong Z et al (2010) Method of parking index based on parking function and location conditions. China J Highw Transp 23(1):112–114

Research on Multimodal Transportation Path Optimization with Time Window Based on Ant Colony Algorithm in Low Carbon Background

Dongxin Yao and Zhishuo Liu

Abstract Green transportation has always been the focus of international attention. With the proposal of "One Belt And One Road" and the emphasis on logistics efficiency and cost at home and abroad, multimodal transport, as an advanced form of transport organization, has been developing continuously. However, there are few studies on carbon emission of multimodal transport in domestic and foreign literatures. In recent years, high-speed railway has become an indispensable way for Chinese tourists to travel, and the use of high-speed railway in the field of logistics also arises at the historic moment. Order is proposed in this paper for a batch of goods, by air, high-speed rail and highway combination of three kinds of transport mode, build the satisfying path capacity constraints, hard time window to minimize the total transportation cost under the restriction of the mathematical model of the total cost including transportation cost of transport costs, transport costs and carbon emissions, and using ant colony algorithm to solve the model, then use different local optimization strategy for processing, finally it is concluded that the optimization results are verified through the calculation example.

Keywords Green transportation · Path optimization · Ant colony algorithm · Multimodal transport · Mathematical model

D. Yao (✉) · Z. Liu
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China
e-mail: 17120773@bjtu.edu.cn

Z. Liu
e-mail: liuzhishuokobe@163.com
URL: http://faculty.bjtu.edu.cn/trans/7984.html

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_77

1 Introduction

With the development of the economy and the constant emphasis on green energy conservation and efficiency, the traditional transportation industry, one of the basic industries of economic development, has been unable to meet the growing social needs as an advanced, energy-saving and efficient transportation. In the way, multimodal transport is very mature in developed countries, and it is still in a developing state in China. Therefore, it is of great significance for the research of multimodal transport and the development of low-carbon traffic. Multimodal transport is a transport mode that uses two or more (at least two) different modes of transport (rail, road, aviation, etc.) to achieve efficient movement of goods. This model can fully exploit the advantages of different modes of transport and achieve "door to door" "Transportation service." It needs to ensure the orderly, smooth and efficient transportation process through reasonable planning, scientific organization and management to achieve the overall optimal goal of the system. There are some different terminologies, i.e., multimodal, intermodal, co-modal and synchro-modal, for the multi-mode transportation systems in the current literature and global logistics industry. Actually, multimodal transportation is the broadest term and covers all of the others. In detail, multimodal transportation is able to provide more efficient, reliable, flexible and sustainable way of freight transportation [1]. Compared with the single transportation mode, multimodal transportation has significant advantages: first, it can effectively reduce transportation costs and improve transportation resource utilization; second, reduce energy consumption and reduce environmental pollution; third, achieve international transportation and domestic transportation. The effective connection is conducive to the development of foreign trade; fourth, shorten the transportation time and accelerate the turnover of funds.

The domestic multimodal transport route optimization literature is generally divided into two categories. The first major category is the known N points, and the goods are transported from the 1st point to the Nth point according to the node order. There are a variety of transportation modes between the points. When the entry and output of the same node adopt different transportation modes, it is necessary to transfer and reload inside the node. For example, the literature [2–6] is to study the multi-modal transport route optimization problem of traversing all nodes. Jing [2] studied the transportation mode selection and multi-objective optimization problem based on hybrid genetic algorithm; Zhang [3] will solve the problem. It is transformed into the shortest path problem with time constraint and capacity constraint and solved by Dijkstra algorithm. Wang [4] studied the combination mode optimization problem of multimodal transport network transportation mode based on Dijkstra algorithm; Luo [5] considered carbon emission in the paper Problem; Lu [6] added time constraints based on previous research. The second type is N nodes, and the nodes are not connected in sequence. The goods are transported from point 1 to point n, but it is not necessary to traverse all points. For example, TONG Lu's multimodal path optimization model and method research, the shortcoming of this paper is that the text only describes the second type of problem., but returned to the first

category when building mathematical models [7]. Liu [8] has studied the comprehensive transportation optimization model of various modes of transportation. Although it does not traverse all nodes, the paper divides the transportation city into supply and demand, and does not actually optimize the second type of problem. Although ZHANG Run-jie's research on multi-modal path optimization based on improved genetic algorithm is also the second type of problem, the paper only establishes a simple mathematical model, without examples and data [9]. ZHANG Ming's paper studied the multi-modal transport problem of 10 nodes through genetic algorithm and Dijkstra algorithm, but the overall traversal situation is similar to the first major category [10]. Ye [11] studied the urban network of 9 nodes and solved it with the ant colony algorithm. The network scale is not very applied to the algorithm. LIAO Tian-bo's paper "Research on Joint Transportation Portfolio Optimization Based on Tabu Search Algorithm" establishes the logistics distribution between specific cities. It is not between any two cities, the model is simple and the result is simple [12].

Representative studies on multimodal transport in foreign countries include: Garcia [13] introduce the advantages of OR technology and AI search methods, and solve the multimodal transport problem by combining the two methods, using automatic planning to achieve each service. The best mode of transport or routing for different pick-up and delivery requests, the solution is currently used in large Spanish logistics companies; Bhattacharya [14] a strategic transport planning model involving a network-wide multimodal transport system for railways/Road intermodal network optimization. The model has been successfully applied to the existing FMCG distribution network in India; Kazemi and Szmerekovsky [15] evaluated the efficiency of a multimodal transport system consisting of pipelines, barges, railways and trucks; Chang [16] proposed Multimodal transport problem. The intermodal routing problem is a multi-objective multimodal multicommodity flow problem (MMMFP) with time windows and concave costs. And a heuristic solution based on Lagrangian relaxation and decomposition techniques is used to decompose the original problem into smaller. The subproblem is then solved.

Through the above literature research at home and abroad, it is found that most of the multimodal transport path optimization problems do not have a feasible low-carbon optimization for complex large-scale networks. This paper fully considers transportation costs, transshipment costs, carbon emission costs, and time window requirements. The combination of aviation, high-speed rail and highway transportation modes, the mathematical model is constructed to minimize the total transportation cost, and the ant colony algorithm is used to solve the problem, which provides an effective solution for the multi-modal transport network path and transportation mode combination optimization problem.

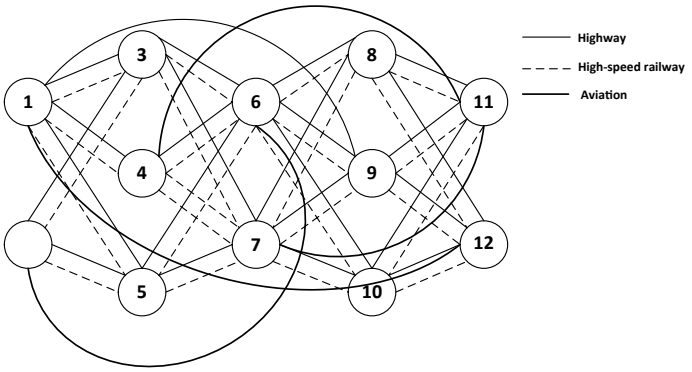


Fig. 1 Small logistics network

2 Mixed Integer Programming Model for Multimodal Transport

2.1 Problem Description

There is a logistics network with a shipment of goods from the starting point to the destination. There are g transportation modes available between any two adjacent cities. The transportation time, cost and transportation capacity constraints between the two adjacent cities are different. When changing from one transportation mode to another, it must meet certain conditions. Transfer time and transfer fee. Under the conditions of hard time window constraints and path capacity constraints, the order route and mode of transport (aviation/high-speed rail/highway) are obtained, so that the total cost of all orders is the lowest. Figure 1 shows a small logistics network.

2.2 Related Assumptions

(1) The transshipment of goods can only occur at city nodes, and each city node can only be reprinted once.
(2) The goods cannot be divided during transportation.

2.3 Mathematical Model

① The symbol description is shown in Table 1.
② Decision variables

Table 1 Symbol description table

Symbol	Significance
Q	Order demand
V_k	Carrying capacity in k modes of transport
T_i^{km}	The transit time required at the node i from the k mode of transport to the m mode of transport
$T_{i,j}^{k}$	Transportation time required from node i to node j in the k mode of transport
C_{ij}^{k}	The unit cost required from node i to node j in the k mode of transport
C_i^{km}	The unit cost of the transfer required from the k transportation mode to the m transportation mode at node i
$X_{\mathrm{ik,im}}$	The order is transferred from the kth transport mode to the m transport mode at node i
$W_{ik,jk}$	From node i to node j in the k mode of transport
$\mathbf{M}_{ik,jk}$	Path capacity from node i to node j in the k transport mode
d_{ij}^{k}	Distance of the kth transport mode from node i to node j
ω	Unit carbon tax
$\mathbf{q}_{ij}$	i to j carrying capacity
$\mathbf{E_k}$	Unit carbon displacement in the k mode of transport
T_n	Require the earliest arrival time
T_p	Require the latest arrival time
K	Collection of all modes of transport
N	Collection of all nodes

$$X_{\mathrm{ik,im}} = \begin{cases} 1, \text{ the k transport mode changes to} \\ \quad \text{the m transport mode at node i} \quad , (\mathrm{n} \neq \mathrm{m}) \\ 0, \text{ else} \end{cases} \tag{1}$$

$$W_{ik,jk} = \begin{cases} 1, \text{ transport between node i and node j} \\ \quad \text{in the kth mode of transport} \qquad , (i \neq j) \\ 0, \text{ else} \end{cases} \tag{2}$$

③ Objective function

$$\sum_{i \in N, k \in K} \left[\mathrm{f}\left(\frac{W_{ik,jk} \times Q_n}{V_k} \right) \times C_{ij}^{k1} \times d_{ij}^{k} \right] \tag{3}$$

$$\sum_{i \in N, k \in K, m \in K} \left(X_{ik,im} \times Q_n \times C_i^{km} \right) \tag{4}$$

$$\omega \left(\sum_{i \in N} \sum_{j \in N} \sum_{k \in K} \mathrm{q}_{ij} \times \mathrm{E}_k \times d_{ij}^{k} \times W_{ik,jk} \right) \tag{5}$$

$$\mathrm{MIN} = (3) + (4) + (5) \tag{6}$$

④ Conditional constraint

(1) Time constraint

$$\left(\sum_{m} X_{ik,im}^{n} + \sum_{j} W_{ik,jk}^{n}\right) = \left(\sum_{m} X_{im,ik}^{n} + \sum_{j} W_{jk,ik}^{n}\right) \tag{7}$$

$$\left(\sum_{m} X_{ik,im}^{n} + \sum_{j} W_{ik,jk}^{n}\right) \leq 1 \tag{8}$$

$$\left(\sum_{m} X_{im,ik}^{n} + \sum_{j} W_{jk,ik}^{n}\right) \leq 1 \tag{9}$$

$$\sum_{m}\sum_{n}\sum_{j}\left(X_{ik,im}^{n} \times T_{i}^{km}\right) \tag{10}$$

$$\sum_{k}\sum_{i}\sum_{j}\left(W_{ik,jk}^{n} \times T_{i,j}^{k}\right) \tag{11}$$

$$T_n \leq (10) + (11) \leq T_p \tag{12}$$

(2) Capacity constraint

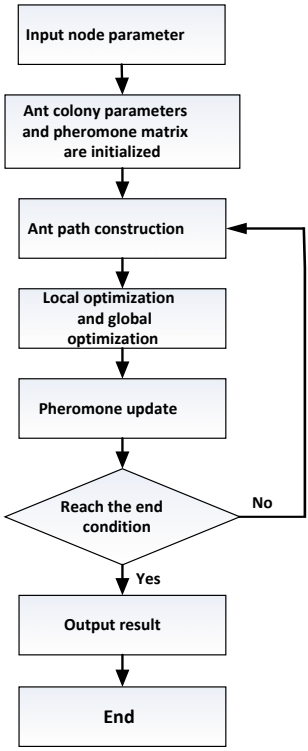
$$\mathrm{Q} < \mathrm{M}_{im,jm} \tag{13}$$

Equation (3) indicates that the node i to the node j are transported by the kth transportation mode; the Formula (4) represents the transit transportation cost at the node i from the kth transportation mode to the m transportation mode; the Formula (5) indicates carbon Emissions cost; Formula (6) represents the objective function; constraint (7) indicates that node i is shipped out of the goods equal to the incoming goods; constraint (8) indicates that node i is only transited or shipped to node j in the shipment; (9) indicates that node i is only transited or transported by node j to i in the shipment; (10) indicates transit time; (11) indicates transit time; (12) indicates hard time window constraint; (13) Indicates path capacity constraints;

3 Multimodal Transportation Algorithm

Ant colony algorithm is widely used in scheduling and optimization problems because of its self-organization, multi-thread parallelism and strong robustness of positive feedback. In this paper, the ant colony algorithm is used to solve the model.

Fig. 2 Algorithm flowchart



The algorithm flow chart is shown in Fig. 2. Step 1: Enter the node parameters; Step 2: Initialize the ant colony parameters and the pheromone matrix; Step 3: Use the ants to generate m routes. Step 4: Perform local optimization and global optimization. Step 5: pheromone update; step six: repeat steps 3–5 until the termination criteria are reached.

For the ant path construction process in the third step of the algorithm, a detailed description is made: at the initial moment, all ants are placed at the initial node of the order ($t_0 = 0$), and the visited customer set $tabu_k = \emptyset$, ant's The driving route routk $= \{0\}$ is located at Lk $= 0$. Ant s as the goods order, starting from the starting point of this order, subject to the order time constraint, path capacity constraints, select some city nodes to reach the end point through appropriate transportation, and thus complete a solution. The second ant repeats the above action until all the ants have finished running, and the end of one iteration ends. The iterations are repeated until the number of iterations is completed and the transport route and mode of transport that ultimately minimizes the cost of order shipping are obtained.

When ant k selects the next node, remove the nodes in $\overline{tabu_k}$ that meet the following constraints and form set $allowed_k$, (1) path capacity constraint: path capacity is $M_{im,jm}$, determined After the route ij satisfies the hard constraint of the path capacity, it is necessary to find a route from the j to the destination D that satisfies the path

capacity constraint: all the paths that do not satisfy the capacity constraint after the j point are excluded, and the path after the exclusion is calculated by the Dijkstra algorithm to the order n. The shortest path of the terminating node, if the path exists, the j point is met; (2) the time constraint: the transit time between the points can be obtained from the parameters of the previous input node, and the ant k calculates the elapsed time t_{oi} every time a node i calculates j Point-to-end point time t_{jD} (excluding the path that does not satisfy the capacity constraint, the shortest time to calculate the end point of the order by the Dijkstra algorithm is t_t_{jD}), if $T_P \leq t_{oi} + t_{jD} \leq T_P$, then j point is met; (3) Calculate the probability P_{ij}^k of each node in the $allowed_k$:

$$P_{ij}^k = \begin{cases} \frac{\tau(i,j)^\alpha \eta(i,j)^\beta}{\sum_{j \in allowed_k} \tau(i,j)^\alpha \eta(i,j)^\beta}, & j \in allowed_k \\ 0, & else \end{cases}$$

$\tau(\mathrm{i}, \mathrm{j})$ is the trajectory strength of the arc (i, j), which is represented by its pheromone. The initial value is 1. After a little time, when the ant completes a cycle, the trajectory intensity changes due to the volatilization of the pheromone. $\tau_{t+1}(i, j) = \rho\tau_t(i, j) + \Delta\tau_t(i, j)$, ρ indicates the pheromone volatilization coefficient, $\Delta\tau_t(i, j)$ denotes the pheromone added in the t cycle, $\eta(i, j)$ is the visibility of the arc (i, j), represented by the cost reciprocal and the time reciprocal; α denotes the relative importance of $\tau_t(i, j)$; β represents the relative importance of $\eta(i, j)$.

4 Instance Design and Result Analysis

4.1 Instance1

Figure 3 shows the logistics network of 35 city nodes. The routes that can be selected

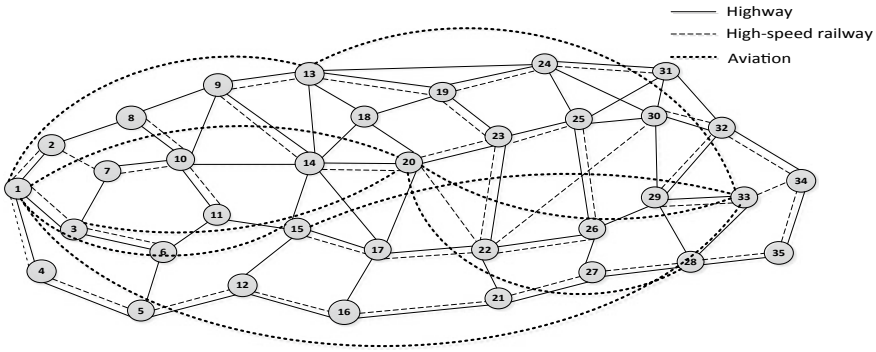


Fig. 3 Instance1

between the city nodes are specified. There are three modes of transportation: highway, high-speed rail and aviation. The transportation modes are available on each route. Also known. When the goods enter the node and the outbound node needs to change the transportation mode, it will generate a certain time and cost. Now it is required to select the appropriate path and transportation mode at the specified time, and transport the 3 tons of goods from the node 1 at the minimum cost. 34, the time window is 12–24 h. Tables 2, 3, 4, 5, 6, 7 and 8 is the necessary data for the instance1. (ω carbon tax price: 100 ¥/t).

4.2 Instance2

Figure 4 shows the logistics network of 48 city nodes. The distance and capacity constraints of different modes of transportation between nodes are not listed because of limited space. The cost, speed, and transshipment cost and time of different modes of transportation are the same as in Eq. 1. The transport task is to transport 6 tons of cargo from node 1 to node 48 with a time window of 12–48 h.

4.3 Instance3

The third example is a large logistics network of 82 city nodes. The distances of different transportation modes between nodes are not listed because of the limited space. The cost, speed, and transshipment cost and time of different modes of transportation are the same as in Eq. 1. The transport task is to transport 6 tons of cargo from node 60 to node 67 with a time window of 12–54 h (Fig. 5).

4.4 Result Analysis

The program is written in Java language and runs on i7 processor, 16 GB memory computer. The parameters are: $\alpha = 1$, $\beta = 1$, $\rho = 0.8$, respectively, no strategy, elite ant strategy, maximum and minimum ant strategy and Remove Strategy The three examples are solved separately, and the following table shows the transportation route and transportation mode of the order, as shown in Table 9.

The optimal transportation route and transportation mode of the three optimization strategies are shown in Table 10 (path: 1, 2, 3… represents the node number, trans: 1/2/3 represents the aviation/high-speed rail/highway respectively).

Compare the results of the three instances using the optimization strategy with the non-strategic results. As shown in Table 11, the brackets are the percentages that are optimized after using the strategy.

Table 2 Transportation distance by different modes of transport (unit: km)

Node	High way	High-speed railway	Aviation	Node	High way	High-speed railway	Aviation	Node	High way	High-speed railway	Aviation
(1,2)	300	250	–	(12,15)	550	510	–	(21,27)	490	430	–
(1,3)	400	380	–	(12,16)	540	500	–	(22,23)	1050	1030	–
(1,4)	730	700	–	(13,14)	770	–	–	(22,26)	540	560	–
(1,13)	–	–	1350	(13,18)	360	–	–	(22,30)	1610	–	–
(1,15)	–	–	1400	(13,19)	670	680	–	(23,25)	430	400	–
(1,20)	–	–	1500	(13,24)	1280	–	–	(24,25)	420	–	–
(1,33)	–	–	3550	(13,33)	–	–	2370	(24,30)	690	–	–
(2,7)	270	–	–	(14,15)	500	–	–	(24,31)	580	550	–
(2,8)	390	–	–	(14,17)	830	–	–	(25,26)	1050	1000	–
(3,6)	440	410	–	(14,18)	410	–	–	(25,30)	350	–	–
(3,7)	450	–	–	(14,20)	430	460	–	(26,27)	300	–	–
(3,20)	–	–	1630	(15,17)	370	390	–	(26,29)	320	–	–
(4,5)	590	560	–	(15,33)	–	–	2480	(27,28)	450	390	–
(5,6)	430	–	–	(16,17)	490	–	–	(28,29)	530	470	–
(5,12)	540	510	–	(16,21)	780	810	–	(28,33)	580	510	–
(6,11)	320	–	–	(17,20)	800	–	–	(28,35)	370	320	–
(7,10)	310	280	–	(17,22)	520	490		(29,30)	670	–	–
(8,9)	450	–	–	(18,19)	380	360	–	(29,32)	690	620	–
(8,10)	330	350	–	(18,20)	400	–	–	(29,33)	370	320	–
(9,10)	650	–	–	(19,23)	420	380	–	(30,31)	250	200	–
(9,13)	430	420	–	(19,24)	530	520	–	(30,32)	280	250	–

(continued)

Table 2 (continued)

Node	High way	High-speed railway	Aviation	Node	High way	High-speed railway	Aviation	Node	High way	High-speed railway	Aviation
(9,14)	840	820	–	(20,22)	880	–	–	(32,34)	580	540	–
(10,11)	430	410	–	(20,23)	470	430	–	(33,34)	220	–	–
(10,14)	620	–	–	(20,33)	–	–	1440	(34,35)	590	540	–
(11,15)	330	–	–	(21,22)	310	–	–				

Table 3 Transportation cost of different transportation modes (unit: ¥/km)

Mode of transport	High way	High-speed railway	Aviation
Cost	4	6	10

Table 4 Transportation cost of different transportation modes (unit: km/h)

Mode of transport	High way	High-speed railway	Aviation
Speed	80	240	800

Table 5 Transit costs between different modes of transport (unit: ¥)

Mode of transport	High way	High-speed railway	Aviation
High way	0	40	60
High-speed railway	40	0	100
aviation	60	100	0

Table 6 Transit time (unit: h)

Mode of transport	High way	High-speed railway	Aviation
High way	0	1	2
High-speed railway	1	0	2
aviation	2	2	0

It can be found from Tables 9, 10 to 11 that the ant colony algorithm can solve this multi-modal path optimization problem well, and the transportation cost and carbon emission cost can be further optimized through different optimization strategies. It can be obtained from Table 11 to Fig. 6. The more nodes have more effective effects in different strategies using heuristic algorithms, the optimization effect of Max-Min strategy is relatively general, and the optimization effect of All Strategies is

Table 7 Carbon emissions from different modes of transport (unit: kg/t * km)

Carbon emission	High way	High-speed railway	Aviation
	0.56	0.14	0.93

Table 8 Carbon emissions from transit transport (unit: kg/t * km)

Carbon emission	High way—high-speed railway	High way to aviation	High-speed railway to aviation
	1.42	2.35	3.45

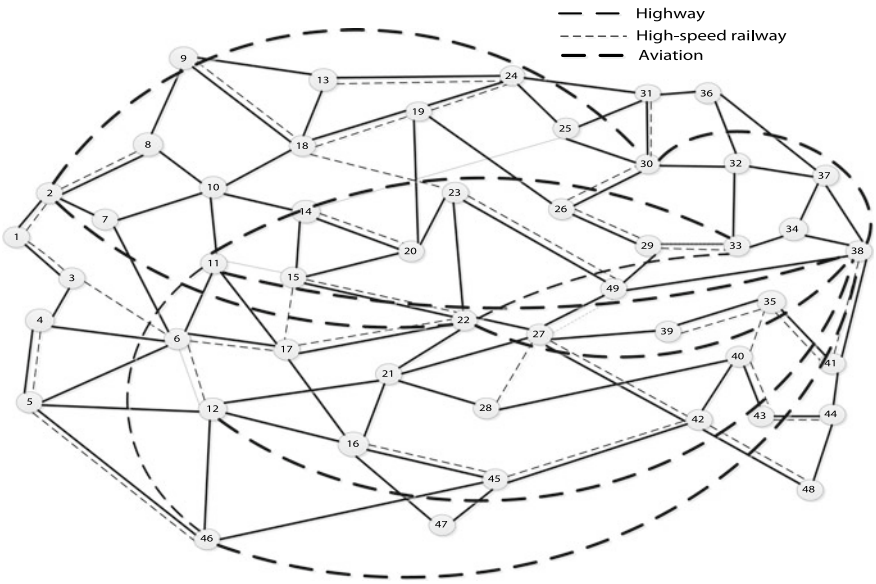


Fig. 4 Instance2

best. Figure 7 shows the iterative convergence of three examples in the case of all Strategies.

5 Conclusion

This paper constructs a mathematical model that satisfies the path capacity constraint, the hard time window constraint, and minimizes the total transportation cost considering the carbon emission cost, and uses the ant colony algorithm to solve the model, and then uses different strategies and local optimization to optimize the process. The optimization results are obtained by numerical examples. The results show that the model and algorithm can effectively provide optimization solutions. In the future research, it can be considered to use multiple logistics orders in the logistics network to carry out transportation services at the same time, which can better reflect the necessity of dynamic change of path capacity constraints.

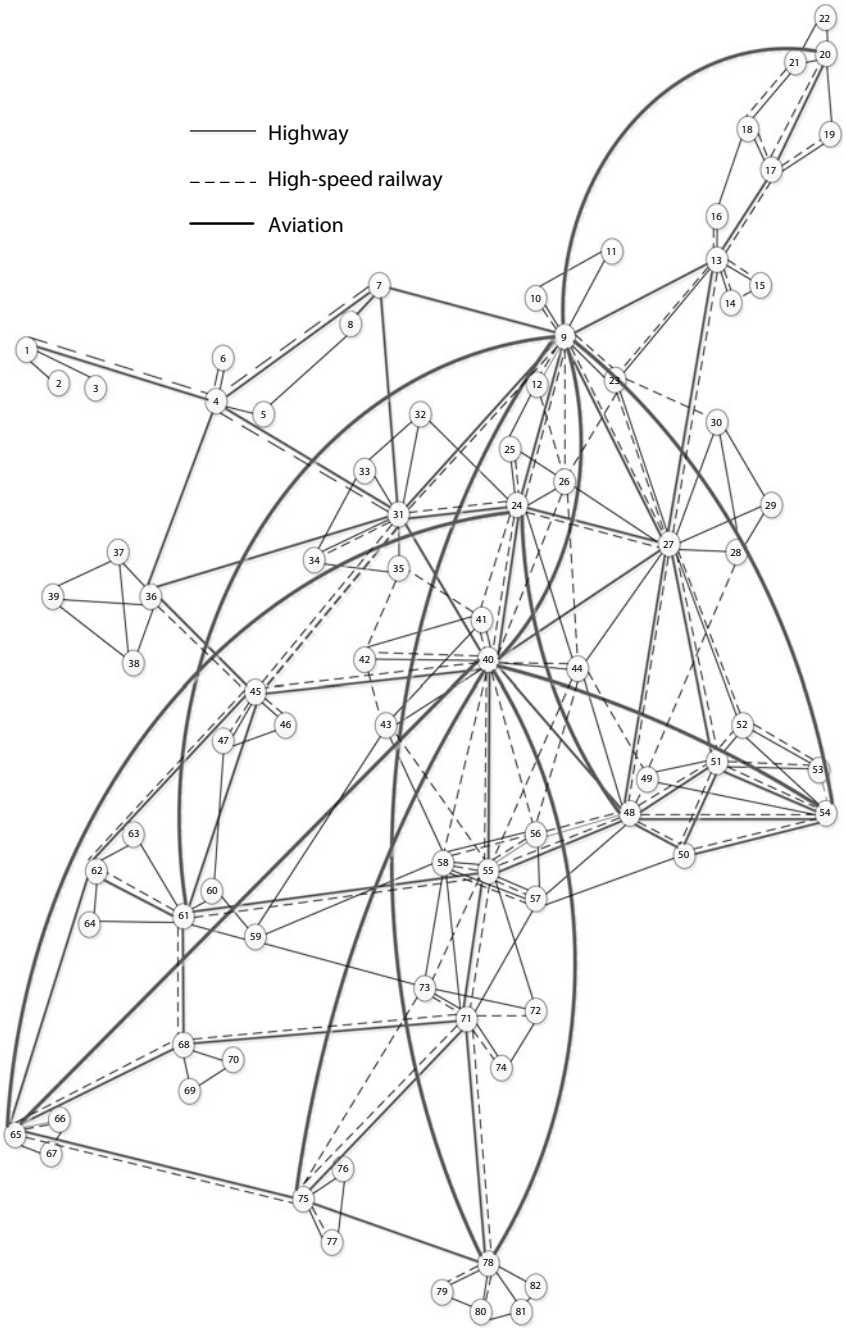


Fig. 5 Instance3

Table 9 The result of algorithm (unit: ¥)

Instance	Optimization strategy									
	None	Time (s)	Elitist strategy	Time (s)	Max-Min	Time (s)	Remove & elitist	Time (s)	All strategies	Time (s)
35 nodes	**27,454**	3.8	27,454	3.4	27,454	4	**26,114**	3	**26,114**	3.2
	27,482	3.8	27,482	3.4	27,482	4	26,114	3	26,114	3.1
	27,482	3.8	**27,454**	3.4	27,454	4	26,114	3	26,114	3.2
	27,482	3.8	27,482	3.4	**27,454**	4	26,114	3	26,114	3.1
	27,454	3.8	27,454	3.4	27,454	4	26,114	3	26,114	3.1
Average–35	27,470		27,465		27,495		26,114		26,114	
48 nodes	12,184	7.2	13,133	7	14,735	7	12,184	6.8	8286	7
	13,133	7	13,133	6.8	14,735	7	8286	6.8	12,184	7
	12,184	7	**8286**	6.5	13,133	7.2	**8185**	6.5	**8286**	6.5
	12,184	7.2	12,184	6.5	13,133	7	13,133	6.8	8286	6.8
	12,184	7	14,735	6.5	**12,184**	7.5	8286	6.5	8286	6.5
Average–48	12,373		12,294		13,584		10,014		9065	
82 nodes	25,724	29	**17,714**	25	18,082	29	**17,714**	22	**17,714**	32
	27,033	27	18,082	29	18,016	30	17,714	22	17,714	30 s
	27,033	27	18,486	29	18,129	30	17,714	22	17,714	32 s
	18,723	27	17,714	24	17,861	31	17,714	22	17,714	32 s
	25,724	27	17,714	23	**17,732**	30	17,714	22	17,714	32 s
Average–82	24,874		17,942		17,964		17,714		17,714	

Bold indicates the optimal result for each optimization strategy for each instance

Table 10 Transportation route and mode of transportation

Optimization strategy	Instance		
	35 Nodes	48 nodes	82 nodes
None	Path: 1 ≫ 4 ≫ 5 ≫ 12 ≫ 16 ≫ 21 ≫ 27 ≫ 28 ≫ 33 ≫ 34	Path: 1 ≫ 3 ≫ 4 ≫ 5 ≫ 12 ≫ 16 ≫ 17 ≫ 22 ≫ 27 ≫ 48	Path: 20 ≫ 17 ≫ 13 ≫ 9 ≫ 27 ≫ 48 ≫ 55 ≫ 57 ≫ 71 ≫ 78 ≫ 75
	Trans: ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 3	Trans: ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3	Trans: ≫ 3 ≫ 2 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 2 ≫ 3 ≫ 3 ≫ 3
	Total cost: 27,454 Total time: 22.125	Total cost: 12,184 Total time: 37	Total cost: 18,723 Total time: 55
Elitist strategy	Path: 1 ≫ 4 ≫ 5 ≫ 12 ≫ 16 ≫ 21 ≫ 27 ≫ 28 ≫ 33 ≫ 34	Path: 1 ≫ 3 ≫ 6 ≫ 17 ≫ 22 ≫ 27 ≫ 48	Path: 20 ≫ 17 ≫ 13 ≫ 23 ≫ 27 ≫ 48 ≫ 57 ≫ 71 ≫ 78 ≫ 75
	Trans: ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 2 ≫ 3	Trans: ≫ 3 ≫ 2 ≫ 3 ≫ 3 ≫ 3 ≫ 3	Trans: ≫ 3 ≫ 2 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3
	Total cost: 27,454 Total time: 22.125	Total cost: 8286 Total time: 21	Total cost: 17,714 Total time: 52
All strategies	Path: 1 ≫ 20 ≫ 23 ≫ 25 ≫ 30 ≫ 32 ≫ 34	Path: 1 ≫ 3 ≫ 6 ≫ 17 ≫ 22 ≫ 27 ≫ 48	Path: 20 ≫ 17 ≫ 13 ≫ 23 ≫ 27 ≫ 48 ≫ 57 ≫ 71 ≫ 78 ≫ 75
	Trans: ≫ 1 ≫ 2 ≫ 2 ≫ 3 ≫ 3 ≫ 2	Trans: ≫ 3 ≫ 2 ≫ 3 ≫ 3 ≫ 3 > ≫ 3	Trans: ≫ 3 ≫ 2 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3 ≫ 3
	Total cost: 26,114 Total time: 19	Total cost: 8286 Total time: 21	Total cost: 17,714 Total time: 52

Table 11 Comparison of results between different strategies

	None	Elitist strategy (%)	Max-Min (%)	Remove & elitist (%)	All strategies (%)
35 nodes	27,470 **(0)**	27,465 **(0.02)**	27,459 **(0.04)**	26,114 **(4.94)**	26,114 **(4.94)**
48 nodes	12,373 **(0)**	12,294 **(0.64)**	13,584 **(−0.98)**	10,014 **(19.07)**	9065 **(26.74)**
82 nodes	24,874 **(0)**	17,942 **(27.87)**	17,964 **(27.78)**	17,714 **(28.79)**	17,714 **(28.79)**

Bold indicates how much the result of each optimization strategy for each instance is improved relative to the result of no optimization strategy

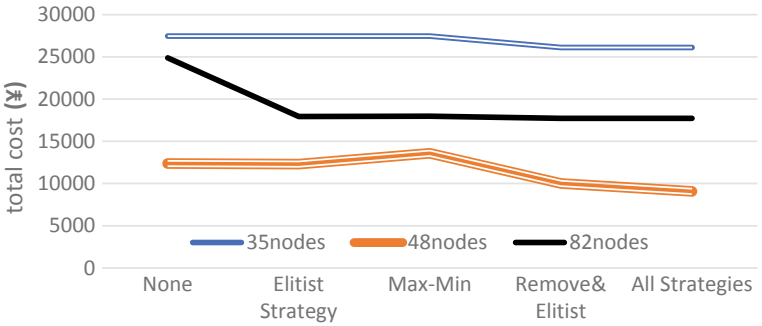


Fig. 6 Comparison of Instances

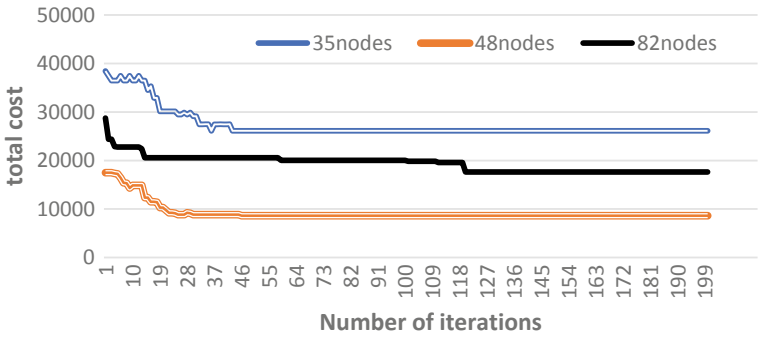


Fig. 7 Iterative graph

Acknowledgements We are grateful for the financial support from the National Key Research and Development Program of China (2017YFB1400100).

References

1. SteadieSeifi M, Dellaert NP, Nuijten W, Van Woensel T, Raoufi R (2014) Multimodal freight transportation planning: a literature review. Eur J Oper Res 233(1):1–15
2. Jing XH, Wei DF, Zhou XZ (2008) Hybrid genetic algorithm for multi-object optimization of transportation modes selection in multimode transportation. Comput Eng Appl 44(6):210–212
3. Zhang DZ, Ling CY (2002) A combination optimization model for multiple transportation mode selection and solution algorithm. J Changsha Railw Univ 04:71–75
4. Tao W, Gang W (2005) A combined optimization mode of multimodal transport network. Eng Sci 10:46–50
5. Luo C (2011) Study on the combination optimization model of multiple transportation modes considering the carbon emission control. J Shaanxi Univ Sci Technol 29(05):113–116
6. Lu X, Lei Q, Wang QC (2012) Study on routing optimization model of inter-modal transportation with time limit. Railw Transp Economy 34(10):52–55
7. Tong T, Nie L, Fu HL (2010) Research on optimization model and method of multi-modal transportation routing. Logis Technol 29(05):57–60
8. Liu SY (2000) Study on comprehensive transport optimization model of multiple transport modes. Transport Sci Technol 05:50–51
9. Zhang RJ (2014) The optimal research on the multi-modal transportation routing based on an improved genetic algorithm. Logis Eng Manage 36(05):95–96
10. Zhang M (2009) Study on the selection of multimodal transportation network path with soft time windows. Central South University
11. Ye HR (2013) Ant colony algorithm for solving transportation problem in the multimodal transportation environment. Hefei university of Technology
12. Liao TB (2015) Research on joint transport combination optimization based on tabu search algorithm. Tsinghua University
13. Garcia J, Florez JE, Torralba A, Borrajo D, Lopez CL, Garcia-Olaya A, Saenz J (2013) Combining linear programming and automated planning to solve intermodal transportation problems. Eur J Oper Res 227(1):216–226
14. Bhattacharya A, Kumar SA, Tiwari MK, Talluri S (2014) An intermodal freight transport system for optimal supply chain logistics. Transp Res Part C 38:73–84
15. Kazemi Y, Szmerekovsky J (2015) Modeling downstream petroleum supply chain: the importance of multi-mode transportation to strategic planning. Transp Res Part E 83:111–125
16. Chang TS (2008) Best routes selection in international intermodal networks. Comput Oper Res 35:2877–2891

RBFNN-Bagging-Model-Based Study on Bus Speed Predication

Xiaoguang Wang, Hai-hua Han, Jin-hui Qie, Si-yang Li, Chun Zhang and Hong-yu Wang

Abstract To establish intelligent bus information systems for the purpose of providing information support for the "smart city" construction, the speed of buses running in the urban road network must be accurately predicted. Common prediction models on bus speed by adopting neural network or supporting technologies like Support Vector Regression (SVR) can well predict vehicle speed on uni-structural sections, but when the prediction scope is extended to the general urban road network (with coexistence of various complex section structures), these models can hardly achieve satisfactory generalization effect, and may generate significant differences in prediction accuracy on different section structures. Therefore, this paper puts forward a RBFNN (Radial Basis Function Neural Network)-based Bagging integrated learning prediction model which can effectively deal with issues concerning the accurate predication of bus speed in the context of general road network. Major research contributions of this paper include: (1) Introducing speed of taxi with sufficient data and a high road coverage rate as the secondary data source so as to make up for sparseness of bus positioning data; (2) Selecting RBFNN as the base model and based on integrated learning philosophy, improving it to RBFNN-Bagging model, which can overcome the shortcomings of uni-structural model and better adapt to differences in section structures. The model raised in this paper, through verification of measured data, has realized an over-90% prediction accuracy rate of bus speed in different sections within the general urban road network, and has witnessed an over-10% promotion in prediction accuracy when compared with that of the neural network and SVR model.

Keywords Bus speed · RBFNN · Bagging · Predication

X. Wang · J. Qie · S. Li · C. Zhang · H. Wang (✉)
China Transport Telecommunications & Information Center (CTTIC),
Beijing 100102, China
e-mail: bitwanghy@163.com

H. Han
College of Information Science and Technology, Yanching Institute of Technology,
Langfang 065201, Hebei, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_78

1 Introduction

1.1 Research Background and Status Quo

With the introduction of "smart city" theory, the transportation sector has providing more and more services oriented towards decision-making purpose, monitoring purpose and travel planning purpose [1], and these services relay heavily on the accurate prediction of road speed. To realize accurate predication of bus speed specific to bus service sector in general transit network can support information services like bus travel planning, bus location query, bus-arrival-time prediction and real-time vehicle dispatch, which are of vital importance to the improvement of transit business and transport service quality [2].

The average speed of vehicle on road section at a specific time is affected by factors such as the historical average vehicle speed on this section, section topology (including characteristics like the length and the number of lanes) and weather condition, among which the historical average vehicle speed on the section can usually be calculated from actual traffic speed of vehicles which have passed through this section. It is necessary to establish a mapping model between numerous influencing factors and the average bus speed for the sake of predicting bus speed. Currently, models that have achieved extensive application and favorable prediction results in this field include nonparametric models like the neural network and Support Vector Regression (SVR), which can carry out accurate mapping on complicated nonlinear relationships. For example, Qiu et al. [3] set up the Bayesian Regularized Neural Network (BRNN) model for urban trunk roads, whose prediction effects can reach above 90% in trunk roads with sufficient data; Asif et al. [4] has expanded the research scope, conducted large-scale road speed predictions over a region covering 5000 road chains, and advanced building a SVM (Support Vector Machine)-based SVR model for each chain, which can achieve favorable overall prediction results.

However, the application of models above has still showed many deficiencies in practice. For instance, practical application verifies that BRNN model failed to achieve ideal prediction effects in some circumstances: (1) the prediction accuracy was less than 60% in sections like the suburbs. The main reason lies in that bus shifts on these sections were few, and sometimes there was even no bus passing by, so the model lacks ample data of bus speed to calculate the accurate historical speed of the predicted sections, which thus influenced the accuracy of speed prediction. (2) Poor generalization performance of the model. For sections in different structures, its prediction accuracy is likely to be fluctuant, which is due to that neural network model is an optimization algorithm for local search, which may easily fall into the local supreme value on calculation of the global threshold (e.g. the prediction results may become multiple-valued when the traffic flow on predicted roads changes frequently), while neural network can hardly well handle such situations, and its prediction effects will become less credible. In view of afore-mentioned deficiencies, this paper has rendered insights for improvement of existing research from the perspectives of data source and mapping methods.

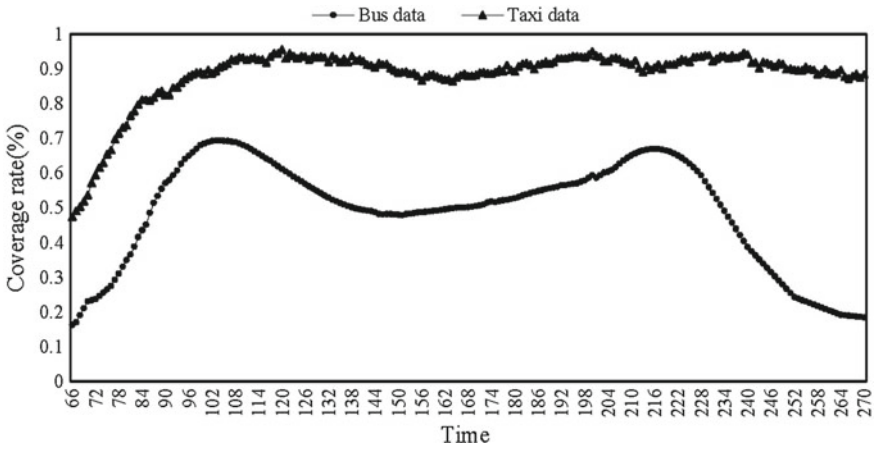


Fig. 1 Variation curve of road coverage rate of buses and taxis

1.2 Optimization of Data Sources

Currently, the prediction of bus speed is mainly based on the data source from bus GPS driving tracks, but due to insufficient bus data and low road coverage, it becomes difficult to achieve accurate prediction of bus speed. For example, there are altogether six million odd motor vehicles but only 12,000 vehicles are equipped with GPS device, accounting for nearly 0.2% of all registered vehicles. Moreover, buses' driving routines like pulling in and receiving passengers have led to its road coverage rate of less than 60% (as shown in Fig. 1). Severe data deficiency can hardly support predictions of bus speed on roads within the general road network. Thus, other data sources should be resorted to for optimization.

The development of vehicle network technology has allowed more registered vehicles to be included into GPS monitoring scope. For example, there are up to 66,000 taxis equipped with GPS device in Beijing, accounting for 1.1% of the total motor vehicles, with a road coverage of over 85%, which goes far beyond that of buses. Considering that there is a strong correlation between bus and taxi speed on the same section, if taxi driving data can be combined with bus driving data and jointly applied to predication on bus driving speed, the accuracy of bus speed prediction will be greatly promoted.

1.3 Optimization of Prediction Model

Different types of prediction models have different characteristics and can be applied to different scenarios. Due to the poor generalization ability of single prediction model which results in difficulties in realizing high prediction accuracy on different

types of sections within the whole transit network, the combination of multiple models can be considered to remove deficiencies in poor generalization ability of single model, so as to attain better overall prediction effects [3, 4].

Inspired by the idea of integrated learning, this paper adopts the Bagging integrated learning framework, and conducted driving speed prediction of buses in general road network on the basis of bus and taxi data. Seeing that RBFNN boasts of better mapping capability and higher computational efficiency than other neural network algorithms [5] and that SVR algorithm has higher computational complexity and lower mapping capability than the neural network [5], RBFNN is chosen as the basis function of Bagging. The number of nodes in hidden layer of RBFNN is adjusted to enhance the diversity of base model and then integrate multiple RBFNN models for the sake of removing single model's sensitivity to changes in input data and enhancing the overall prediction effect.

Based on the above analysis, this paper optimizes the existing prediction methods of bus driving speed from the perspectives of data source and model, and the main contents are as follows: (1) Introducing data source of taxis with a higher coverage rate, so as to deal with relatively low prediction accuracy caused by bus data volume on suburban roads; (2) Proposing a BRFNN-based Bagging integrated learning algorithm to deal with poor generalization ability of single model and difficulties in realizing accurate prediction of the general road network.

2 Application of Multi-source Floating Car Data for Prediction

2.1 Analysis of Floating Car Data's Influences on Calculation of Road Speed Information

Floating Car refers to GPS-device-equipped traffic-information sampling vehicles which can upload vehicle positioning information in real time. The amount of floating cars has a significant impact on the accuracy of traffic calculation. Turner and Holdener [6] have studied on the relationships between parameters like traffic information service level, information credibility, information release cycle and the number of floating cars, and offered Eq. 1 to calculate the amount of floating cars required by urban traffic information services in large cities.

$$n = \left(z^2 \cdot c^2\right)/e^2 \tag{1}$$

In Eq. 1, n is the number of floating cars required, z is the coefficient corresponding to confidence, c is a certain road chain's coefficient reflecting changes in vehicle speed in sampling time interval, and e is the fault-tolerant coefficient. The results of this study indicate that, at 90% level of confidence and 10% fault-tolerant rate, there should be driving date of 2–3 floating cars on each chain every five minutes so as to

fulfil the accurate calculation of traffic information. However, the amount of buses in real driving cycle on many bus lines usually cannot meet this requirement.

In a city, the number of taxis is often more than that of buses. In Beijing, for example, road coverage rate of bus driving data is less than 60% most of the time, while that of taxi data is higher than 80% under the same conditions and even surpasses 90% between 7 am and 23 pm. As shown in Fig. 1.

In Fig. 1, the horizontal axis takes 5 min as the interval time, and the vertical axis is the ratio of the amount of roads covering data of more than 3 buses and taxis in each time period to the total amount of roads in transit network.

As can be seen from the figure, to introduce taxi data source with sufficient data will become a good supplement in calculation of bus speed on roads.

However, due to evident differences between bus and taxi in maximum speed, acceleration, driving routes and other driving rules, if taxis' GPS data is directly used in calculating bus speed, a larger deviation in results will be produced. Therefore, this paper will discuss the feasibility of fusion calculation of taxi and bus data source in the next chapter.

2.2 Fusion of Taxi Data and Bus Data

Although bus and taxi's driving characteristics vary, their speeds on the same section will be affected by traffic density. As shown in Fig. 2, at a low traffic density, vehicles run freely, and buses can run smoothly in this case, which thus brings about simple prediction. However, due to sparseness of bus driving data, taxi data is required to judge the traffic condition; as the traffic density increases, differences in their speed gradually decline and the correlation becomes stronger. In this case, despite

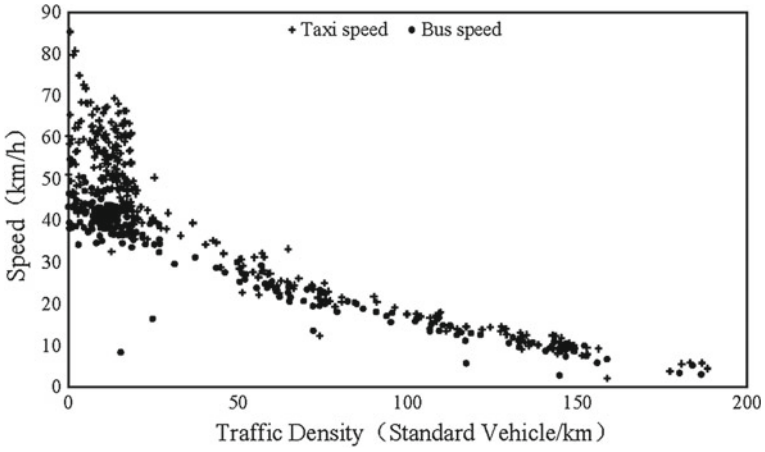


Fig. 2 Correlation analysis between bus speed and taxi speed

Table 1 Prediction results of bus speed by neural network

Data source	Chain 1 (%)	Chain 2 (%)	Chain 3 (%)
Bus data	77	75	63
Bus, taxi data	88	85	79

the relatively complex traffic conditions, difficulties in prediction will elevate, but taxi driving speed can accurately judge the status of buses, so taxi data can provide support for bus speed prediction. Then, taxi data's effects on bus speed prediction will be verified through experimental data.

It can be seen from survey in the first chapter that, neural network is relatively effective in studying traffic information prediction. With specific urban trunk roads as the research object, Park et al. [7] applied neural network model to prediction of travelling time of single urban trunk road, and this method is also applicable to exploration and prediction on variation laws of speed on expressway as well as trunk roads in other cities. Hence, in order to test taxi speed's contribution to prediction of bus speed, this paper randomly selects two chains (chain 1 and 2) among closed roads on North Third Ring Road in Beijing as well as a chain (chain 3) on the relatively complex Zhichun Road, and conducts predictions on bus speed respectively through bus data alone as well as bus and taxi data on the basis of neural network model. The prediction results are shown in Table 1.

It can be indicated from Table 1 that: with added taxi data, prediction accuracy on closed roads increases by 10% or more; while on the complexly-structured Chain 3, even if the accuracy fails to reach 80%, it grows by 16%. Thus, it follows that taxi data can well supplement bus data.

However, compared with prediction results of three chains in the above figure, it can be seen that despite the promotion brought by added taxi data, the prediction results are still not satisfying, so this method lacks feasibility for large-scale prediction. This paper will then improve the prediction model to realize more accurate prediction on bus speed in transit network.

3 Analysis of Input Parameters in Model

3.1 Analysis of Input Parameters in Model

- Parameters of Road's Physical Structure

As the vehicle carrier, roads' physical structure has a decisive impact on vehicle speed. Physical factors that usually affect road speed primarily include the number of lanes, road width, road length, road width at the entrance/exit and the narrowest place. Zhao et al. [8] proposed that major factors influencing urban road speed also include transit stops, whose influences on road speed are mainly manifested in the negative influences of weaving lane-switching movement of buses in entering and

leaving the stops on road speed, and main influencing factors here include bus stops' distance from road entrance/exit, namely the length of weaving area in front of and behind the stop. Therefore, this paper will collect data of road-structure-related parameters, which include road length, road width at the narrowest point, the number of lanes, transit stops and so forth.

- Traffic Cycle

The morning and evening peak is the break point of changes in urban traffic flow. Therefore, in morning and evening peak hours as well as non-working days, significant fluctuations emerge in road speed, so this paper has built the adjustment matrix A of average speed in a long period of time by fitting Eq. 2:

$$Standard\ Devition_{(date,\,time)} = \bar{v}_{(date,\,time)} - \alpha \cdot \bar{v} \tag{2}$$

In Eq. 2, date parameter $date = \{holiday, weekend, weekday\}$, time period parameter $time = \{peek,\ flat\ peak,\ idle\ hours\}$. $Standard\ Devition_{(date,\,time)}$ corresponds to the speed adjustment factor of data and time; $\bar{v}_{(date,\,time)}$ represents the average speed corresponding to data and time, $\bar{v}$ represents the overall historical average speed, and the unit of the above three parameters is m/s; α represents the adjustment factor.

$$A = \begin{bmatrix} 1.01 & 0.94 & 0.57 \\ 1.29 & 1.26 & 1.18 \\ 1.41 & 1.44 & 1.49 \end{bmatrix}$$

Rows in matrix A respectively represent holiday, weekend and weekday, while columns there respectively represent peak, flat peak and idle hours from the top down. It has expressed adjustment value of average speed in holiday, weekend and weekday during peak, flat peak and idle hours. According to regulations over the number of vehicles on road as well as the morning and evening peak in Beijing, the morning peak period lasts from 7:00 to 9:00 while the evening peak period from 17:00 to 19:00, and other periods are flat peak. Due to influences of various factors, the beginning and end time of the morning and evening peak are not fixed. This paper will revise the division of morning and evening peak via reference to IBEE-based serial sample clustering [9], which has predicted the break point of speed during morning and evening peak through speed changes in road chain in a day.

Numerous studies have shown that road speed will be affected in different weather conditions to different extents. Smith et al. [10] studied on road speed and traffic capacity under different rainfalls, finding that sprinkle will decrease traffic capacity by 4–10% and speed by 5–6.5%; a heavy rain will decrease traffic capacity by 25–35% and speed by 5–6.5%. In order to facilitate calculation, the factor weather is divided into four grades and then converted into numerical value.

Average Road Speed in the First N Time Periods Under normal circumstances, the speed of road will generally not change frequently, and the vehicle driving interval of urban buses is usually 5 min, so 5 min is selected as the interval for dividing time

periods. Since road traffic flow usually changes within a cycle of half of hour, $N = 6$, and the average road speed in the first 6 time periods is selected as the input.

3.2 Model Construction

In order to accurately predict bus speed and serve intelligent traffic applications like the calculation of bus arrival time, based on dynamic driving data of floating cars and weather data on urban roads, this paper constructs an optimized integrated learning algorithm to predict bus speed.

3.2.1 Selection of Methods

Compared with single model, integrated learning has unparalleled advantages in terms of accuracy, generalization ability and local optimization. Besides, since the prediction of urban road speed is a complex multi-classification prediction program, this paper adopts integrated learning for modelling.

For integrated learning, there are two main development directions: Bagging and Boosting [11]. In contrast, Bagging is more suitable for a multi-classification prediction program, and trains every classifier in parallel, so it has a natural advantage in parallel processing of big data [12, 13]. A successful variant algorithm of Bagging is the Random Forest (RF) [14]. Although RF has added the random selection of attributes and made great progress when compared with Bagging, it fails to take the importance of various factors into consideration. In such a time-series application scenario like road speed prediction, recent data can provide more information than long-dated data, and the weight between these attributes is different. If Bootstrap is still adopted to obtain the training subset, it will result in the loss of data information, which will thus affect the effect of prediction. Secondly, the decision tree is a weak classifier, and can hardly achieve satisfactory results for complex mapping. Therefore, this paper conducts optimization in terms of diversity enhancement and combination of base model, and utilizes spark's distributed parallel processing idea to construct a model suitable for real-time speed prediction of urban roads, thereby efficiently and accurately carrying out real-time bus speed prediction.

3.2.2 General Procedures of Bagging

1. For original training sample set, adopt Bootstrap for sampling and obtain N training set (s).
2. Select the base model, conduct parallel training of base model in each training set, and obtain N classifier (s).
3. The *N* classifiers are integrated to obtain an integrated classifier.

This paper will conduct optimization in respect of diversity enhancement:

- Diversity Enhancement

The key to integrated learning is to enhance the diversity among learners, which is the so-called "harmony in diversity". This part first analyzes the importance of model diversity for model accuracy, and then proposes a strategy to elevate the algorithm accuracy by increasing the diversity between various models.

- Diversity Analysis

Assuming that the base learner set is $\{h_1, h_2, \ldots, h_T\}$, and the learner H is integrated through simple weighted average. Its equation is as follows:

$$H(x) = \sum_{i=1}^{T} \omega_i \cdot h_i(x) \tag{3}$$

Calculate the integrated model H and the base model h_i in the regression learning task $R^d \rightarrow R$. For x, define the difference between two learners as $D(h_i|x)$.

$$D(h_i|x) = (h_i(x) - H(x))^2 \tag{4}$$

Then there is the total error $\overline{D}(h_i|x)$

$$\overline{D}(h|x) = \sum_{i=1}^{T} \omega_i \cdot D(h_i|x) = \sum_{i=1}^{T} \omega_i \cdot (h_i(x) - H(x))^2 \tag{5}$$

Obviously, the difference term obtained here shows the diversity of the base learner in sample set. The square error of base learner h_i and integrated learner H is respectively:

$$E(h_i|x) = (f(x) - h_i(x))^2 \tag{6}$$

$$E(H|x) = (f(x) - H(x))^2 \tag{7}$$

Let $\overline{E}(h|x) = \sum_{i=1}^{T} \omega_i \cdot E(h_i|x)$ denote the weighted mean of the base learner's error, so there is:

$$\overline{D}(h|x) = \sum_{i=1}^{T} \omega_i \cdot E(h_i|x) - E(H|x) = \overline{E}(h|x) - \overline{E}(H|x) \tag{8}$$

Let $\overline{E} = \sum_{i=1}^{T} \omega_i \cdot E(h_i|x)$ denote the error mean of the base learner on the data set while $\overline{D} = \sum_{i=1}^{T} \omega_i \cdot D(h_i|x)$ denote the distance mean between the base learner and the integrated learner, so there is:

$$E(H|x) = \overline{E} - \overline{D} \tag{9}$$

Equation 9 clearly indicates that given that constant prediction accuracy of base learner, the greater the difference between learners is, the better the integration effect will be.

- Diversity Enhancement Strategy

This paper will strengthen the diversity from the aspects of the number of hidden layer nodes of RBFNN and the integration of models.

(1) The Number of Hidden Layer Nodes

The number of hidden layer nodes has always been the discussion focus of researchers on RBF neural network, and the number n_i is determined by the strategy of Fisher decision criteria proposed by Huang et al. [5]. Since the sample training set obtained by Bootstrap sampling is not consistent with the distribution of the original data set, it is necessary to adjust the number of hidden layer nodes in different data sets, which can not only enhance the diversity of model, but also make the base model more suitable for its sampled data and thus improve the accuracy of the integrated model more effectively.

$$m_i = \frac{v_{i,\max} - v_{i,\min}}{k_i} \tag{10}$$

$$n_{i,new} = \begin{cases} n_i - \sqrt{n_i - m_i} & m_i \leq n_i \\ n_i + \sqrt{m_i - n_i} & m_i > n_i \end{cases} \tag{11}$$

As the variance of each training sample's labelled data represents the complexity of its prediction, this paper calculates the post-sampling average range of labelled data with intensive samples by Eq. 10, and adjusts the number of hidden layer nodes by Eq. 11.

(2) Integrated Strategy of Base Model

At present, the commonly-used classifier integration methods include: simple average, weighted average, super majority vote, relative majority vote (random selection of one result), weighted voting learning method, etc. Each method has its corresponding application scenario. This paper is to adopt a model integration method through determining the weight and using the weighted average according to the prediction accuracy of each base model.

3.3 Implementation of the Optimized Model

- Specific Procedures

(1) Read the training dataset S, and store S in RDD form.

(2) Conduct Bootstrap sampling on RDD, and generate 10 training sets $U = \{U_1, U_2, \ldots, U_{10}\}$.
(3) For the 10 training sets generated, use RBF neural network model for parallel training, and obtain 10 classifiers $F = \{F_1, F_2, \ldots, F_{10}\}$, among which F_i has recorded the parameters of the i-th classifier.
(4) For the 10 classifiers obtained in Step 3, predict the original training data and obtain the prediction result set $R = \{R_1, R_2, \ldots, R_{10}\}$.

(5) According to prediction results and real results of each classifier obtained in Step 4, calculate the error of the classifier. The error calculating equation of the i-th classifier is:

$$E_i = ERROR(F_i) = \frac{\sum_{j=1}^{|S|} S_j - R_{i,j}}{\sum_{j=1}^{|S|} S_j} \tag{12}$$

(6) Assign weights to the classifier according to the error of each classifier. And the weight calculating equation of the i-th classifier is:

$$W_i = 1 - E_i \tag{13}$$

(7) Obtain the ultimate combining classifier FC through comprehensive consideration of results and weights of each classifier, and FC is defined as:

$$FC = \frac{\sum_{i=1}^{|S|} F_i \cdot W_i}{\sum_{i=1}^{|S|} W_i} \tag{14}$$

4 Experiment and Evaluation

Through testing the data set, this paper verifies the accuracy of the proposed road speed prediction model, and illustrates the advantages of the model in this paper by comparing it with other models.

4.1 Test Data Set

Driving data useful in the test include relevant data of 20,000 buses and 66,000 taxis in recent two years. 85% of the data get involved in the training of model, while other data serves as test data of the model.

As the main driving area for buses is confined to the downtown area, according to the Urban Road Grade Proportion [15], this paper divides roads into expressway, trunk road and secondary trunk road. For expressway, with fewer signal lights, vehicles will run at a high speed. And for trunk road, due to influences of signal lights' interference, vehicles may start and stop frequently, and factors like the diversity of road structure will make the driving speed more diversified. For secondary trunk road, factors like mixed traffic flow and interlacing trunk and side roads, changes in driving speed will become all the more complicated. Such a division of road structure can fully compare the exact influences of testing non-linear single models and integrated models in different application scenarios.

4.2 Evaluation Criteria of Model

The commonly-sued evaluation criteria adopted in this paper for predicting model: the average relative error. It represents the average size of overall prediction error, and the calculating equation of the average relative error δ_{MAPE} is:

$$\delta_{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\overline{V}_{prediction} - \overline{V}_{actual}\right|}{\overline{V}_{actual}} \tag{15}$$

Here is the accuracy calculating equation:

$$Accu = (1 - \delta_{MAPE}) \cdot 100\% = \left(1 - \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\overline{V}_{prediction} - \overline{V}_{actual}\right|}{\overline{V}_{actual}}\right) \cdot 100\% \tag{16}$$

In Eqs. 15 and 16, $\overline{V}_{prediction}$ is the predicted value, $\overline{V}_{actual}$ is the actual value.

4.3 Experimental Scheme

In order to fully compare the nonlinear single model and the optimized integrated model in this paper, BRNN [3] model and SVR [4] model, which have presently shown favorable road speed prediction effect, are selected and tested on identical test data. Secondly, since the size of floating cars' data volume becomes a bottleneck

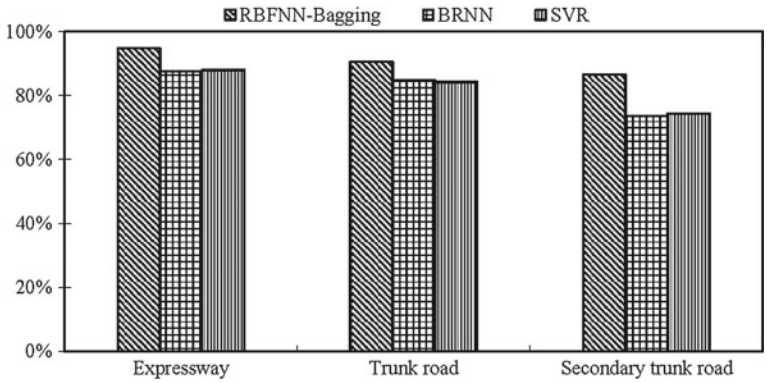


Fig. 3 Comparison of model accuracy in different

influencing the prediction accuracy of road speed, this paper divides roads into three categories (sufficient, moderate, few) according to the size of bus data volume, and compares BRNN with SVR model through utilizing taxi data as supplementary test data so as to explain that model in this paper can well accomplish prediction on bus speed on roads in the case of multi-source data. Finally, this paper verifies the feasibility of this model by means of the measured data.

- Adaptability to Different Types of Roads

As traffic speed is closely related to road structure and traffic operation laws in urban Beijing, this paper will test three types of roads of different structures in downtown area, including expressway, trunk road and secondary trunk road. The test results are shown in Fig. 3.

It can be seen from Fig. 3 that BRNN's accuracy of bus speed prediction on expressways and trunk roads only remains at 88 and 84% or so, and fails to perform better, which is mainly due to BRNN model's strong sensitivity to data and poor generation ability. So it performs poorly in prediction on some roads, thereby leading to a decline of BRRN model's accuracy in the entire road network. Moreover, on account of the complex structure of secondary trunk roads, BRNN model performs poorly in secondary trunk roads, with an accuracy of less than 75%. Although SVR model's accuracy performance resembles that of BRNN model, it is less efficient and less applicable to actual conditions.

RBFNN-Bagging model in this paper reveals quite stable prediction effects in different road structures, and the accuracy has been maintained at 85% and above. From the comparison between models, accuracy of this model is 8–15% higher than that of other models, especially on trunk roads and secondary trunk roads. The generalization ability of this model allows it to better adapt to the cases of complex traffic and sparse bus data.

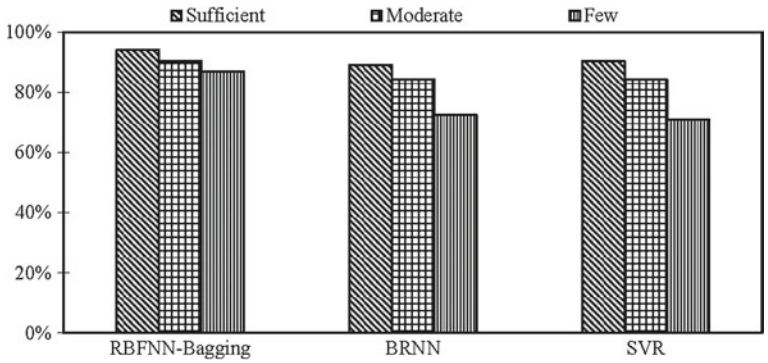


Fig. 4 Model contrast under different data volumes

- Adaptability to Different Data Volumes

The shortage of floating car data remains as a difficult problem in road speed prediction. The size of data volume becomes a restriction on research scope and depth of intelligent transportation. This paper will divide transit network in Beijing into three grades according to bus driving data volume: sufficient (more than 2 vehicles per 5 min on average), relatively sufficient (more than 1 vehicle per 10 min on average) and insufficient (1 or less than 1 vehicle per 15 min on average). It calculates the accuracy of each model in each data volume grade, and the results are shown in Fig. 4.

It can be seen from Fig. 4 that these models vary little under the condition of sufficient data volume. As data volume decreases, the advantage from this model's strong generalization ability becomes gradually distinct, and the model can even maintain fine stability under conditions of sufficient/insufficient data volume, and differs within 5% from the prediction accuracy in the case of sufficient data volume; while other two models show poor stability, and their accuracy also decreases by 15% or so. It is proved that compared with BRNN and SVR model, model in this paper boasts of stronger adaptability. Integrated learning can well overcome deficiencies of single model, and smoothly predict bus speed within general road network under the condition of uneven distribution of buses.

- Validation of Measured Data

This paper predicts bus speed on a road chain in general road network within 80 min, and compares that with actually-collected bus speed data. The comparison results in Fig. 5 have intuitively showed the accuracy of model in this paper, thereby demonstrating that it can well predict the bus speed in a future period of time.

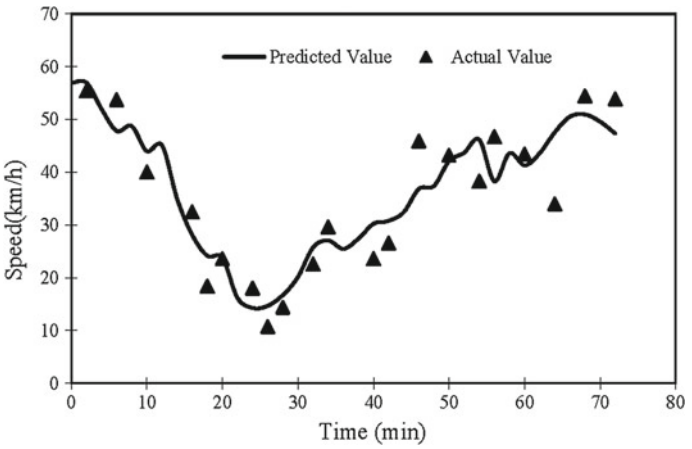


Fig. 5 Comparison between model prediction and measured data

5 Conclusion

In this paper, we propose a RBFNN-Bagging model to predict bus speed in a future period of time. This model initiatively selects two data sources, namely bus and taxi, as the input data for the purpose of compensating for the low prediction accuracy caused by sparseness of bus data volume and the resulting failure to acquire some input parameters like the average speed of historical roads. In addition, taking the influences of weather and static road attributes like road width and the number of lanes on vehicle speed into account, this paper collects physical parameters of weather and road, analyzes their correlation with road speed, and determines input parameters of the model.

Owing to the extensiveness of prediction scope and complexity of input data, based on ideas of integrated learning, this paper starts from the selection of base model and its diversity enhancement, and puts forward the RBFNN-Bagging model with RBF neural network as the base model, which can carry out preferable prediction on speed variation on complex roads. In the experiment, this paper compares the prediction effects of BRNN, SVR model and model in this paper on different types of roads, and analyzes and contrasts prediction accuracy of each model under the condition of different bus data volumes. The experimental results demonstrate that model in this paper boasts of strong generalization ability, and can render reasonable prediction, with no need for further correction, in general road network and roads with low bus data volume.

However, since model input in this paper contains parameters like local weather and bus operation laws, this model may have certain dependence on locality. Thus, we will further conduct comparative study on data of different regions, so as to determine their influences on the model.

References

1. Hancke GP, Silva BDCE, Hancke GP (2012) The role of advanced sensing in smart cities. Sensors 13(1):393–425
2. Huang E, Antoniou C, Wen Y, Ben-Akiva M, Lopes J, Bento J (2009) Real-time multi-sensor multi-source network data fusion using dynamic traffic assignment models. In: International IEEE conference on intelligent transportation systems
3. Qiu C, Wang C, Zuo X, Fang B (2011) A bayesian regularized neural network approach to short-term traffic speed prediction. In: IEEE international conference on systems, man and cybernetics, Anchorage, Alaska, USA, Oct, pp 2215–2220
4. Asif MT, Dauwels J, Goh CY, Oran A (2014) Spatiotemporal patterns in large-scale traffic speed prediction. IEEE Trans Intell Transp Syst 15(2):794–804
5. Huang J, Wang Y, Liu Z, Guan B, Long D, Du X (2016) On modeling microscopic vehicle fuel consumption using radial basis function neural network. Soft Comput 20(7):2771–2779
6. Turner SM, Holdener DJ (1995) Probe vehicle sample sizes for real-time information: the Houston experience. In: The Pacific Rim TransTech Conference. 1995 Vehicle navigation and information systems conference proceedings. 6th international VNIS. A ride into the future, pp 3–10
7. Park J, Li D, Murphey YL, Kristinsson J (2011) Real time vehicle speed prediction using a neural network traffic model. In: International joint conference on neural networks, pp 2991–2996
8. Zhao N, Yu L, Zhao H, Guo J, Wen H (2009) Analysis of traffic flow characteristics on ring road expressways in beijing using floating car data and remote traffic microwave sensor data. Transport Res Rec J Transport Res Board 2124(2124):178–185
9. Sun Z, Wang Y, Pan J (2008) Short-term traffic flow forecasting based on clustering and feature selection. IEEE, pp 577–583
10. Smith BL, Byrne KG, Copperman RB, Hennessy SM, Goodall NJ (2004) An investigation into the impact of rainfall on freeway traffic flow. In: 83rd annual meeting of the Transportation Research Board, Citeseer, Washington, DC
11. Woniak M, Manuel A, Corchado E (2014) A survey of multiple classifier systems as hybrid systems. Inf Fus 16(1):3–17
12. Islam MM, Yao X, Shahriar Nirjon SM, Islam MA (2008) Bagging and boosting negatively correlated neural networks. IEEE Trans Syst Man Cyber Part B Cyber A Publ IEEE Systems Man Cyber Soc 38(3):771–784
13. Wang H, Jin L, Hou Z, Fang R, Mei W, Jian H (2017) Research on parallelized real-time map matching algorithm for massive GPS data. Clust Comput 20(2):1123–1134
14. Kulkarni VY, Sinha PK (2013) Random forest classifiers: a survey and future research directions. Int J Adv Comput 36(1):1144–1153
15. Juan HE, Wei D, Qizhou HU, Liu Z (2008) Urban road grade proportion based on trip mode. J Transport Syst Eng Inf Technol 8(5):88–93

Research on Applying Solar Energy Technology to Rail Transit Vehicle

Yanwei Lu, Wang Xing, Jialin Zhou, Mintang Sun, Shufeng Li and Xinying Hou

Abstract Rail transit vehicle consumes a great deal of power in operation, while applying solar energy technology could reduce the consumption of electric energy. This paper researches on the solar energy technology applying on rail transit vehicle's hot water supply and photovoltaic power generation system. Hot water supply system of solar energy which can supply hot water through day and night consists of solar collector, water tank, molten salt heat exchange system and control system. Photovoltaic power generation system connects in parallel with rail transit vehicle's charger to supply electric energy to storage battery and DC load, it consists of buck-boost chopper, photovoltaic cell and power generation controller.

Keywords Solar energy · Photovoltaic cell · Solar collector · Rail transit · Molten salt · Buck-boost chopper · Energy conservation and emission reduction

1 Introduction

"Made in China 2025" proposes to adhere to "innovation-driven, green development" and implement projects such as intelligent manufacturing engineering, green manufacturing engineering and high-end equipment innovation. Meanwhile, environmental friendly new energy rail transportation is one of the development directions in CRRC Intelligent Manufacturing 2025, which is the action outline of CRRC. Research on applying solar energy technology to rail transit vehicle conforms the national strategy, as well as reduces energy consumption and carbon dioxide emissions, therefore reducing the expenses of train operating. Solar energy is a kind of green energy, which is crucial to promoting rail transit vehicle's innovation, green manufacturing, the innovation and upgrading of advanced rail transit equipment

Y. Lu (✉) · J. Zhou · M. Sun · S. Li · X. Hou
CRRC SIFANG Co., Ltd., No. 9 Hongping Road, Qingdao 266111, Chengyang District, China
e-mail: kphongkong@126.com

W. Xing (✉)
16-6A, West 4th-Ring Mid Road, Beijing 100000, Haidian District, China
e-mail: xw_job@126.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_79

manufacturing industry and the competitive strength of high-speed rail "going out" strategy.

2 Solar Radiant Energy

Sun radiation [1] is originated from the thermonuclear fusion reaction under high temperature and high pressure. During the thermonuclear fusion reaction, four protons become one helion, the mass is lost, while huge energy is released, as shown in the following Formula:

$$4^1\mathrm{H} \underset{\Delta \mathrm{m}}{\rightarrow} {}^4\mathrm{He} + 2\mathrm{e}^+ + \Delta E$$

where $\triangle m$ is the lost mass in reaction, is 4.29×10^{-29} kg. According to Einstein mass-energy relation:

$$E = mc^2$$

where E is energy, c is the speed of light (3×10^8 m/s). Therefore, when the amount of 4.29×10^{-29} kg mass loss occurs, 3.86×10^{-12} J energy is released. Calculating with current thermonuclear reaction rate, the lifetime of sun is 5×10^9 years.

Affected by the existing of the atmosphere, sun radiation that reaches the earth's surface can be defined as direct radiation and scattered radiation. Generally speaking, direct radiation accounts for a large proportion of total radiation in sunny days, while scattered radiation accounts for a large proportion of total radiation in cloudy and rainy days. The available solar power that reaches the earth's land surface is 1.785×10^{16} W. China has abundant solar energy resources, the theoretical reserves amount to 1700 billion tons of standard coal per year, which has great development potential.

3 The Constitution of Rail Transit Vehicle's Solar Energy System

Rail transit vehicle's solar energy system consists of independent solar energy hot water supply system and photovoltaic power generation supply system. Hot water supply system contains flat plate solar collector, molten salt heat exchange system, water supply controller, water tank on the vehicle's roof, electric boiler (see Fig. 1). Photovoltaic power generation system contains photovoltaic cell, rail transit vehicle's storage battery, power generation controller, buck-boost chopper, lightning arrester etc. (see Fig. 2). Hot water supply system is integrated assembled at the end of the vehicle to supply hot water to passengers by water supply pipeline and electric boiler. Photovoltaic cell component is assembled to the middle part of the vehicle's

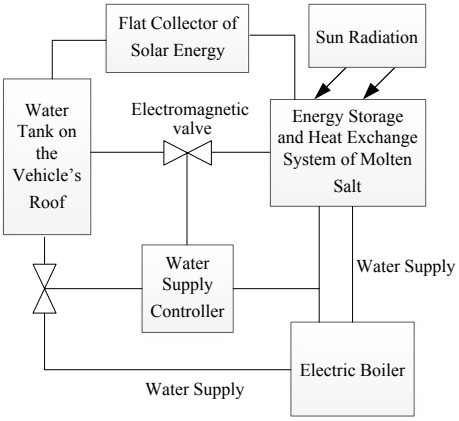


Fig. 1 Block diagram of solar energy hot water supply system

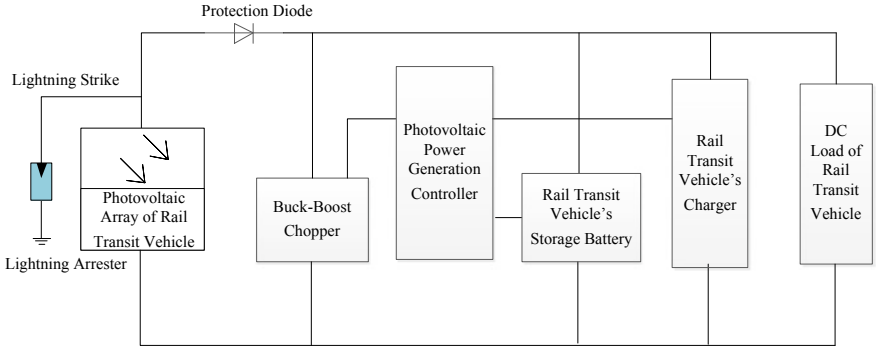


Fig. 2 Block diagram of photovoltaic power generation and supply system

roof, supplying DC power to the vehicle through power generation controller and buck-boost chopper.

3.1 Flat Plate Collector

Flat plate collector is featured with simple structure, reliable operation, large heat absorption area and fixed installation location (it does not have to track the sun). Flat plate collector can make use of direct radiation and scattered radiation at same time, which makes it adapt to rail transit vehicle's operations. Flat plate collector is a special heat exchanger, the fluid in its heat receiver could obtain useful energy income through energy transformation with solar radiant. To decrease the heat loss in the heat conduction of flat plate collector with external environment, insulation

materials are equipped to flat plate collector's back and side. The energy balance equation of flat plate collector is as follows:

$$Q_A = Q_U + Q_L + Q_s$$

where Q_A is the solar radiation energy absorbed by the flat plate collector per unit of time, Q_U is the flat plate collector's usable energy output per unit of time, Q_L is the heat loss of the flat plate collector per unit of time, Q_S is the energy stored by the flat plate collector per unit of time.

3.2 Heat Exchange of Molten Salt

Solar energy storage and heat exchange technology of molten salt utilizes raw materials such as nitrate as heat transfer medium, this technology stores energy through internal energy conversion of solar radiation with molten salt. Energy storage and heat exchange technology of molten salt has high hot melt and heat conduction value, high heat storage density and compact structure of heat storage unit. It also has good thermal stability and mass transfer speed, the heat exchange and water supply capability at night while store energy at daytime as well as relatively low price.

Energy storage [2] technology of molten salt has been applied to solar thermal power station. The principle is heating the molten salt to a molten state by solar radiation energy, the molten salt flow exchanges heat with water through the pipeline, thus producing water vapor steam to generate electricity. Different from the above technology application, the energy storage and heat exchange system of molten salt absorbs the thermal energy from solar radiation energy and flat plate collector, exchanges heat with water flow in the pipeline. With less complexity and low maintenance costs, this system is more applicable to rail transit vehicle.

3.3 Photovoltaic Cell

Photovoltaic cell [1] is the smallest unit of photoelectric conversion. The photovoltaic array consists of photovoltaic cells is assembled on the rail transit vehicle's roof surface as part of the roof. Photovoltaic cells have anti-corrosion, wind-proof, anti-icing and rain-proof capabilities to some extent. Circuit diagram and equivalent circuit diagram of photovoltaic cell under illumination are shown in Figs. 3 and 4, respectively, where R_L is battery load resistance, I_{SC} is short circuit current. The value of I_{SC} is related to the size of photovoltaic cell, I_{SC} is proportional to the size. The value of I_{SC} is 16–30 mA with the size of 1 cm^2 for the photovoltaic cell (the value in the photovoltaic power generation system is 25 mA). I_{SC} increases slightly when ambient temperature increases, I_{SC} increases about 78 μA when every 1 °C increases

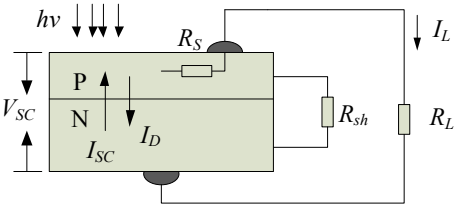


Fig. 3 Circuit diagram of photovoltaic cell under illumination

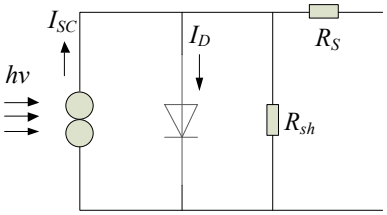


Fig. 4 Equivalent circuit diagram of photovoltaic cell under illumination

in temperature commonly. The voltage of the monocrystalline silicon photovoltaic cell is approximately 450–600 mV (the value in the photovoltaic power generation system is 500 mV),the maximum voltage can be 690 mV. In Figs. 3 and 4, where I_D is the total diffusion current through P-N knot, whose direction is opposite to $I_{SC.}$ R_s is the series resistance consists of battery's volume resistance, surface resistance, electrode conductor resistance, contact resistance of electrode with silicon chips' surface. R_{sh} is the by-pass resistance, caused by unclean edges of silicon chips or internal defect.

R_s of an ideal photovoltaic cell is rather small, while R_{sh} is rather big. As R_s connects in series and R_{sh} connects in parallel in the circuit, their values could be ignored during ideal circuit calculation. The current I_L flows through the load R_L is:

$$I_L = I_{SC} - I_D$$

P-N knot's equation of the characteristic curve is:

$$I_{\mathrm{L}} = I_{\mathrm{SC}} - I_0\left(\mathrm{e}^{\frac{qV}{AkT}}\right)$$

where I_0 is the reverse saturation current of photovoltaic cell without illumination, q is the electronic charge, k is boltzmann constant, A is the curve factor of diode. When $I_{\mathrm{L}} = 0$, the voltage of the photovoltaic cell V_{OC} can be expressed as:

$$V_{\mathrm{OC}} = \frac{AkT}{q}\ln\left(\frac{I_{\mathrm{SC}}}{I_0} + 1\right)$$

According to the above two equations, the current-voltage relation curve of photovoltaic cell can be illustrated (see Fig. 5). Curve (a) illustrates the current voltage characteristic curve of diode in dark, i.e. *I-V* curve without illumination. Curve (b)

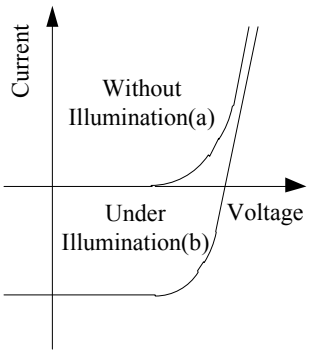


Fig. 5 The current-voltage relation curve of photovoltaic cell

illustrates *I-V* curve of photovoltaic cell under illumination, obtained from curve (a) shift toward the fourth quadrant with a value of I_{SC}.

Available area on rail transit vehicle's roof for photovoltaic cell assembling has a length of 19 m, a width of 3 m. The effective area S = 57 m^2 (570,000 cm^2), therefore the power of usable photovoltaic cell is:

$$P = S \times I_{SC} \times V_{SC} = 570{,}000 \times 0.025 \times 0.5\,kW = 7.125\,kW$$

3.4 Rail Transit Vehicle's Storage Battery

Photovoltaic power generation system supply electricity to emergency load together with storage battery under rail transit vehicle's failure, extending emergency time effectively. In regular operation, photovoltaic power generation system supply electricity to DC load together with rail transit vehicle's charger, using photovoltaic power generation system in preference in order to decrease energy consumption. Extra electric energy can be used to charge the storage battery for emergency. Here the authors use the case of EMU (Electrical Multiple Units) trains to explain the collaboration work of photovoltaic power generation system and storage battery. Technical specifications of the EMU's storage battery application are:

(1) During power outage, storage battery's volume should support auxiliary equipment's operation more than 30 min.
(2) Storage battery should maintain working for more than 2 h as emergency power supply, and be charged through circuit in operation.
(3) When external power shut off, car lighting, wireless communication, broadcast device, front and rear external sign lights and emergency ventilation device etc. should maintain working for more than 2 h only by storage battery's power supply.

Table 1 Photovoltaic power generation system's effect of strengthening the emergency work capability when the train is in failure mode

EMU train's failure mode	Maximum DC load/kW	Train's operation lifetime/h	Storage battery life	The effect of photovoltaic power generation system
EMU train's power failure	6.477	2.0	2 h 55 min	To support train's operation
External power failure	3.104	2.0	6 h 45 min	The train can keep operating under illumination without electric energy from storage battery
Wire power failure or 2 devices of Auxiliary Power Unit (APU) failure	14.412	0.5	60 min	To suspend train's operation lifetime to 2 h

The case of CRH2 EMU train's storage battery [3] can be used to explain photovoltaic power generation system's effect of strengthening the emergency work capability when the train is in failure mode [1] (see Table 1).

3.5 Photovoltaic Power Generation Controller

Photovoltaic power generation controller [4] connects in parallel in photovoltaic power generation system. The switching device T1 connects in parallel at the output terminal of the system (see Fig. 6). Photovoltaic power generation controller collects circuit's voltage and current in real time. When rail transit vehicle's storage battery does not allow charging, the switching device T1 conducts to discharge the output current from photovoltaic power generation system by short circuit. Therefore, storage battery will not be overcharged, T1 has the effect of overcharge protection.

The switching device T2 is the discharge control switch of rail transit vehicle's storage battery. When the current of DC load is greater than rated current, T2 shuts down to protect the system from overload and short circuit. When the voltage of storage battery is less than over-discharge voltage, T2 shuts down to protect the storage battery from over discharge.

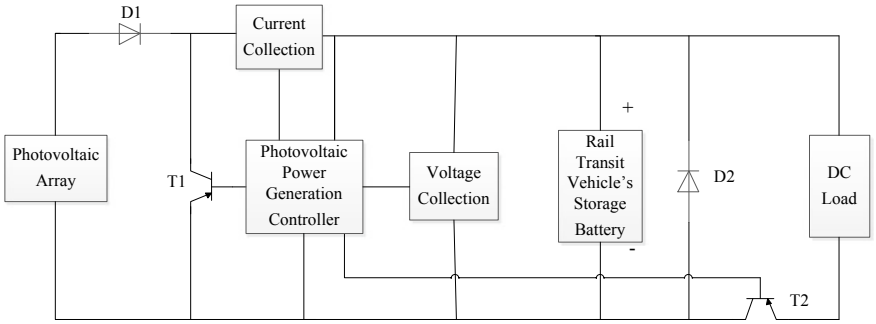


Fig. 6 Control principle of photovoltaic power generation controller

D1 is the anti-reverse charging diode. Only the output voltage of photovoltaic array is greater than the voltage of storage battery can D1 conducts. Therefore, photovoltaic array will not back-discharge at night or in rainy days, D1 has the effect of anti-reverse charging.

3.6 Buck-Boost Chopper

There are variety kinds of rail transit vehicles, so do the power supply voltage classes. The main classes are DC110, DC48 and DC24 V. In order to adapt to different kinds of rail transit vehicles as well as improve the resistant capability of voltage fluctuation, buck-boost chopper is deployed at the output terminal of photovoltaic power generation system [5]. Default output voltage of the system is DC110 V, different voltage demand can be met through setting parameter code of output voltage when adapt to different kind of rail transit vehicle. As shown in Fig. 7, buck-boost chopper circuit consists of switching device T3, inductance L, capacitance C and diode VD. The polarity of the circuit's output voltage u_o is opposite from voltage E, this circuit is an inverting output converter.

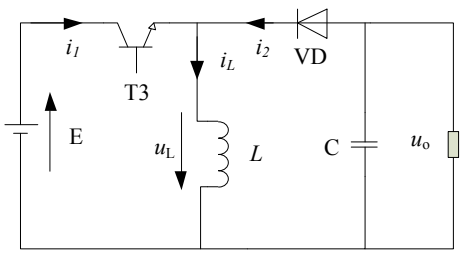


Fig. 7 Circuit diagram of buck-boost chopper

3.7 Arrester

Arrester connects in parallel in the photovoltaic power generation system, it is in the state of non-conducting normally. When overvoltage happens to rail transit vehicle's photovoltaic power generation system because of lightning strike, arrester breaks down and conducts immediately to discharge, cutting overvoltage down and conducting lightning charge to earth, therefore, protecting photovoltaic power generation system from lightning strike. After overvoltage disappears, arrester restores to the normal state of non-conducting. Photovoltaic power generation system utilizes zinc oxide arrester.

4 Operating Principle of Rail Transit Vehicle's Solar Energy System

Rail transit vehicle has a large passenger capacity, the demand for hot water is rather huge especially for long-distance train. Electric boiler would consume a great deal of electric energy to supply hot water, while solar energy hot water supply system even can supply boiled water when illumination is sufficient as well as save electric energy effectively. The energy storage system of molten salt absorbs and store thermal energy under illumination, then makes use of the stored energy to guarantee hot water supply at night by heat exchange system.

Output voltage is controlled by photovoltaic power generation system through power generation controller and buck-boost chopper, matching the voltage of rail transit vehicle's charger and storage battery in order to powering DC load. Rail transit vehicle would use photovoltaic power generation system in preference. The system not only can charge the rail transit vehicle's storage battery when illumination is sufficient, but also supply power to emergency load with storage battery at the time of power supply system failure or power off completely.

4.1 Operating Principle of Solar Energy Hot Water Supply System

Solar energy water tank, heat exchange system of molten salt and light sensor are assembled on the rail transit vehicle's roof, hot water supply is automatically controlled by water supply controller according to sensor's signal. Flat plate collector and solar energy water tank are integrated designed in solar energy hot water supply system, and filter system is utilized to remove impurities such as scale generated by solar energy collector system. This system also has the function of lightning strike protection. Operating principle of solar energy hot water supply system can be shown in Fig. 8.

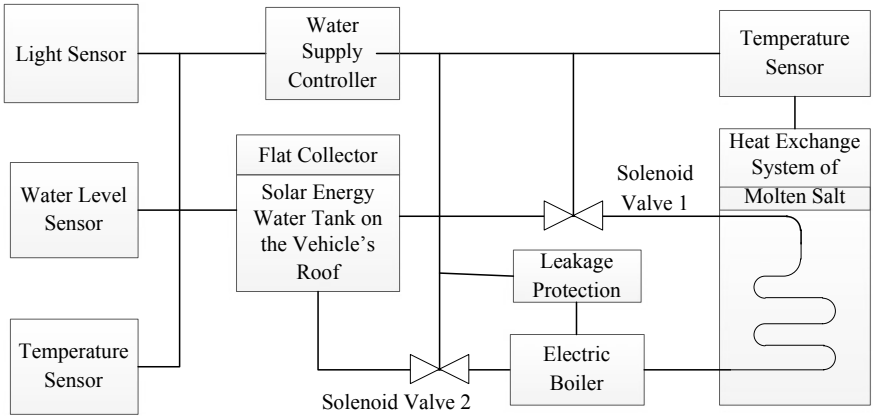


Fig. 8 Operating principle of solar energy hot water supply system

Flat plate collector absorbs solar radiation energy and exchanges heat with water tank. Heat exchange system of molten salt absorbs heat and stores energy under illumination to support heat exchange at night and in rainy days. When the light sensor of the water tank on vehicle's roof detects light signal, water supply controller keeps solenoid valve 1 off while solenoid valve 2 on during electric boiler's operating period. At night or in rainy days, water supply controller turns solenoid valve 2 off and solenoid valve 1 on, the system supplies hot water by heat exchange system. The water flows through heat exchange system of molten salt in pipeline, water temperature increases rapidly and flows into electric boiler.

Water supply controller collects sensor signal from solar energy water tank on the vehicle's roof. When water temperature is above 95 °C, exchange system of molten salt is off. Water supply controller detects the sensor signal from heat exchange system of molten salt and solar energy water tank on the vehicle's roof in real time. Therefore, when the temperature of molten salt and water is the same, heat exchange stops, solenoid valve 1 is off during water supply period. When solar energy water tank and heat exchange system can not supply water above 95 °C, electric boiler operates automatically.

4.2 Operating Principle of Photovoltaic Power Generation System

Photovoltaic power generation system consists of photovoltaic array, power generation controller, buck-boost chopper, rail transit vehicle's charger, storage battery and load. Power generation controller collects the voltage of storage battery's circuit in real time and controls buck-boost chopper's state of on and off. Photovoltaic power generation system could release lightning overvoltage when lightning strike happens,

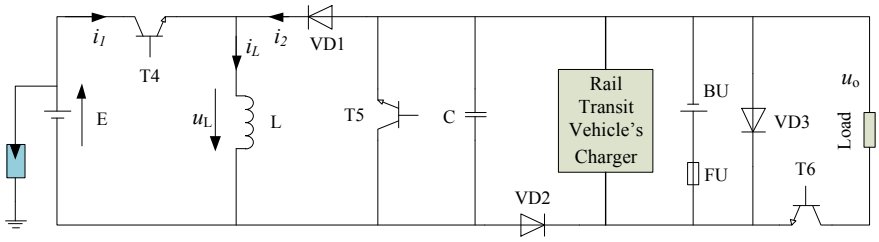


Fig. 9 Operating principle of photovoltaic power generation system

guarding rail transit vehicle's power supply. Figure 9 shows the operating principle of photovoltaic power generation system.

Rated output voltage of rail transit vehicle's photovoltaic power generation system is DC110 V, voltage fluctuation range is DC100 ~ 135 V, buck-boost chopper is utilized to adapt to different DC load, as different rail transit vehicle's voltage class varies. The average value U_0 of buck-boost chopper's output voltage could be greater or less than the voltage of photovoltaic power generation E, their polarity is opposite. When the switching tube T4 conducts, the current passed inductance L becomes greater gradually under the effect of photovoltaic power generation's voltage E. The energy is stored in inductance, diode VD1 is reverse biased and cut off, capacitor C maintains the output voltage U_0 substantially constant and supply power to storage battery and DC load at that time. After that, the switching tube T4 is off, diode VD1 conducts forward under the influence of inductance voltage and output voltage. Energy stored in inductance releases to output circuit through diode VD1, the current is i_2, the direction of current is shown in Fig. 9. Diode VD1 is an anti-reverse charging diode, diode VD3 is an anti-reversal diode of storage battery. Diode VD3 conducts when the polarity of storage battery is reverse, storage battery discharge by VD3 and fuses the fuse FU.

When the photovoltaic power generation system is in steady state, the integral of the voltage u_L across the inductor L is zero in one cycle, it can be shown as follows:

$$\int_0^T u_L \mathrm{d}t = 0$$

When the switching tube T4 of the photovoltaic power generation system's buck-boost chopper conducts and stores energy(conduction time is t_{on}), $u_L = E$ (the voltage of photovoltaic cell). When T4 is off, energy releases to storage battery and DC load (shut off time is t_{off}), $u_L = -u_o$, therefore:

$$Et_{on} = U_0 t_{off}$$

Output voltage of the photovoltaic power generation system is:

$$U_o = \frac{t_{on}}{t_{off}} E = \frac{t_{on}}{T - t_{off}} E = \frac{a}{1 - a} E$$

where α is conduction rate, it can be adjusted to adapt to different assembling conditions of photovoltaic power generation system. When $0 < \alpha < 1/2$, the voltage of photovoltaic power generation system can be chopped to DC48 or DC24 V. When rail transit vehicle's load voltage is DC110 V, α should be set in the interval $1/2 < \alpha < 1$, to prevent low-voltage protection caused by system's direct voltage fluctuation, the photovoltaic power generation system increases voltage.

When the rail transit vehicle operates normally, photovoltaic power generation system's power generation controller sends command to adjust conduction rateα, making output voltage greater than that of rail transit vehicle's charger. At the same time, the voltage of rail transit vehicle's charger decrease accordingly, photovoltaic power generation system supply power to load and storage battery directly. When photovoltaic power generation system failure occurs, rail transit vehicle's charger operates automatically to guarantee power supply. Overvoltage protection is effective as T5 conducts to discharge the output current of photovoltaic power generation system by short circuit. When DC load's current is greater than rated current and lead to overload or short circuit, T6 shuts off to protect the system from overcurrent.

5 Economic Benefits and CO_2 Emission Reduction Analysis

Peak sunshine duration is to convert the amount of local solar radiation to hours under standard test conditions, that is, ratio of total solar radiation H to standard irradiance G on the ground surface (irradiance G = 1000 W/m^2), or equivalent utilization hours. Rail transit vehicles operate throughout China, solar radiation belts exist in such huge operation area. This paper takes peak sunshine duration as 1750 h/year into consideration.

5.1 Economic Benefits

The on-roof water tank's capacity is 400 L = 0.4 m^3, insulation materials are utilized to water tank, pipelines and heat exchange system of molten salt to reduce heat loss. Calculating with peak sunshine duration and rail transit vehicles' operation time 5 h, hot water consumption volume is 400 L, the total energy Q water tank absorbed is as follows (per year):

$$\mathrm{Q} = \mathrm{T}cm\Delta t = 3.7 \times 10^{10}\,\mathrm{J} = 10{,}220\,\mathrm{kWh}$$

where m is water's mass, which is 400 kg; c is water's specific heat, which is 4.2×10^3 J/(kg °C); Δt is the volume of water temperature change, which is 60 °C; T is the day's number of system operation, which is 365.

Design operating temperature is 250 °C, heat exchange system of molten salt insulates from the external by insulation system. For the purpose of saving energy as at daytime, heat exchange volume of heat exchange system per day $Q_{Exchange} = Q/365$ J $= 10^8$ J $= 28$ kWh, the mass of molten salt m is:

$$m = \frac{Q_{Exchange}}{c_{Molten\ Salt} \times \Delta t} = 355\,\text{kg}$$

where m is the mass of molten salt; c is molten salt's specific heat, which is 4.2×10^3 J/(kg °C); Δt is the volume of molten salt temperature change, which is 200 °C.

One car can save energy $Q_{each\ car/year} = 2Q = 20440$ kWh per year by utilizing solar energy hot water supply system. Electricity generated by one car's photovoltaic power generation system $W = PtT = 13003$ kWh per year. Saving energy $W_{year} = Q_{each\ car/year} + W$=33,443 kWh per year by utilizing solar energy technology on one car of rail transit vehicle.

Molten salt materials' price is 526 RMB yuan/kWh, the cost of one car is 15,000 RMB yuan; water tank on the roof is from original design, there is hardly any extra cost; flat plate collector, insulation system and hot water supply controller cost 5000 RMB yuan; the total cost of solar energy hot water supply system is 25,000 RMB yuan for one car. The price of photovoltaic cell is about 5 RMB yuan/W, photovoltaic power generation system would cost 60,000 RMB yuan per car. As the design life of rail transit vehicle is 30 years, one car can save nearly 1 million RMB yuan by utilizing solar energy technology.

5.2 CO_2 Emission Reduction Analysis

Solar energy is a kind of green renewable energy source, CO_2 emission could be reduced because of using solar energy. Emission reduction potential equation of solar energy hot water supply system and photovoltaic power generation is as follows:

$$PM = K \times P_0 \times Y_r \times N \times EI$$

where PM is CO_2 emission reduction potential, the unit is kg; K is the efficiency of synthetic system, is 0.7; P_0 is the power of solar energy system (converted power), is 18 kW; Y_r is peak sunshine duration for one year (added with converted hours at night), is 2100 h; N is the design life of rail transit vehicle, is 30 years; EI is CO_2 emission index, is 0.814 kg/(kWh) in China.

Calculation result shows, one car could reduce 646 ton of CO_2 emission in 30 years.

6 Conclusion

Applying solar energy technology to photovoltaic power generation and hot water supply system on rail transit vehicle, this is a new researching area and has good promotion value and social benefits. Solar energy technology is green and environmental-friendly, which can reduce CO_2 emission and promote the development of the rail transit industry. Solar energy technology also can reduce operating costs and increase income. High speed train is an iconic product of "One Belt One Road", green technology and operation can improve the competitiveness of China's manufacturing. Solar energy hot water supply system and photovoltaic power generation applied to one car can save cost 0.915 million RMB yuan during whole lifetime, which has a bright future of development.

References

1. Li A, Lyu Q (2015) Solar photovoltaic power generation system engineering. The Second Edition, Chemical Industry Press, Beijing, 3, pp 124–129
2. Wang X, Wang L (2013) Application progress of molten salt thermal storage technology in new energy industry. Electr Manuf 10:74–78
3. Hou X, Lu Y (2017) Application of photovoltaic power generation in rail transit vehicle. Urban Mass Transit 6:138–141
4. Wang C, Wang S (2009) Practical technology of solar photovoltaic power generation. The Second Edition Chemical Industry Press, Beijing, pp 13–16
5. Zhaoan W, Jun H (2008) Power electronic technology. The Fourth Edition, Mechanical Industry Press, Beijing, pp 100–111

Yanwei Lu (1986–), Male, Master's Degree, Engineer, working in CRRC SIFANG Co., Ltd., engaged in the design and development of railway vehicle electrical system, traction system calculation, etc. Address: CRRC SIFANG Co., Ltd., No. 9 Hongping Road, Chengyang District, Qingdao, China.

Wang Xing (1989–), Female, Master's Degree, Assistant Engineer, working in CRRC International Co., Ltd., in charge of strategic planning, enterprise operations etc. Address: 16-6A, West 4th-Ring Mid Road, Beijing 100000, Haidian District, China.

Analysis of Typical Attacks on Intelligent and Connected Vehicle Cyber Security

Xingshu Liu and Ling Yang

Abstract The research on risk attacking and vulnerability mining of vehicle cyber security has been carried out all over the world, and the hidden dangers and problems of cyber security exposed by different types of automobiles have become more and more frequent. How to build an ecological circle of cyber security of the whole automobile industry chain still has a long way to go. This paper takes Ford Winged Tiger as the research object, and makes a detailed analysis of the typical cases of the brake system being attacked. By changing the CAN bus data message, the accelerator pedal will fail, which provides a reference for the management and prevention of the cyber security of intelligent and connected vehicle.

Keywords Intelligent and connected vehicle · Cyber security · CAN bus · Attacking

1 Introduction

Electricity, networking, intelligence, and sharing have become the inevitable trend of the development of the automobile industry. With the development of intelligent vehicle networking technology, more and more automobile electronic control systems are available. The automobile will no longer be an isolated unit, but will become a mobile intelligent network terminal. Based on intranet, inter-vehicle network and vehicle cloud network, wireless communication and cyber exchange between vehicle and vehicle, vehicle and road, vehicle and pedestrian, vehicle and the internet are

X. Liu
School of Mechanical Engineering, Tianjin University of Technology and Education, Tianjin 300222, China

Hunan Automotive Engineering Vocational College, Zhuzhou 412001, China

L. Yang (✉)
School of Information Technology Engineering, Tianjin University of Technology and Education, Tianjin 300222, China
e-mail: 335351859@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_80

carried out according to the agreed communication protocol and data exchange standard, and the interconnection system of intelligent traffic management, intelligent dynamic cyber service and intelligent vehicle control is constructed. However, while the interconnection system brings comfort, convenience and efficiency to people, it also brings more and more serious cyber security risks. Vehicles connected by intelligent networks will become "supercomputers" on the road, which store huge amounts of data to provide efficient intelligent services. However, this change also provides opportunities for various security vulnerabilities, such as hackers and viruses on industrial equipment, **which makes the vehicle under the intelligent network face the risk of illegal intrusion and attack**. In the white paper published by the famous white hat hackers Miller and Dr. Valasek in 2013 DEFCON, how to attack the control systems of Toyota Prius and Ford Wing Tiger to achieve abnormal behavior such as brake failure or sudden braking at high speed, and even including the source code, compiler and schematic diagram of the connector used for the attack [1]. At the Black Hat USA 2014, they released a research report on the cyber security of more than 20 models on the market, assessing the ability of different automobile manufacturers to withstand malicious attacks on different models [2]. At the Black Hat Conference in 2015, they demonstrated that they could use a laptop to remotely control a Jeep Cherokee by attacking an on-board entertainment system through a "0 day" vulnerability [3]. These publicly published research reports have pushed the cyber security issues of network vehicle from the edge to the foreground, in order to arouse the attention of automobile manufacturers and researchers on the cyber security issues of network vehicle.

2 Intelligent and Connected Vehicle Threat Model

Cyber security risk of automotive electronic system can be divided into internal risk and external risk, as shown in Fig. 1.

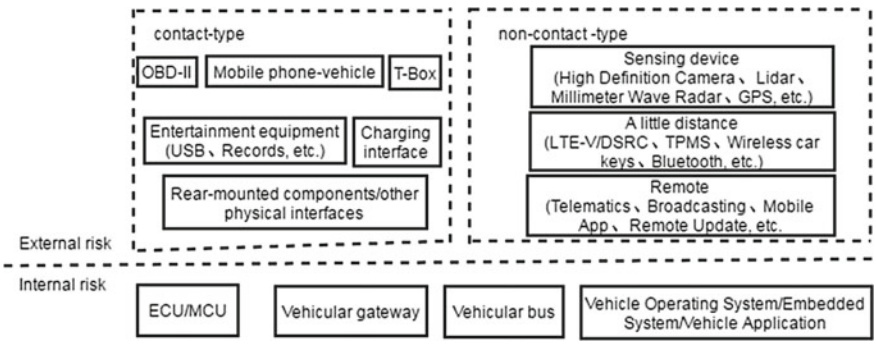


Fig. 1 Cyber security threat model of intelligent and connected vehicle

2.1 Internal Risks

Internal risk refers to the safety hazards that may be caused by intruding into the interior of a vehicle. The main sources of risk are ECU/MCU, bus, gateway, on-board operating system/embedded system/on-board application, and so on.

(1) Currently, many on-board electronic controllers (ECU/MCU) are imported from abroad. Even some domestic controllers are programmed and developed by using foreign chips. Once the back door or loopholes are left in the ECU/MCU, **it may lead to cyber being secretly transmitted or the control of the system being covert**.
(2) Common vehicular buses, including LIN, CAN, FlexRay and MOST, all have international standards and use open bus protocols to communicate. Once intruded into the bus system, it is easy to be injected malicious pseudodata to affect the normal communication, and even to obtain the control of the whole vehicle.
(3) The network topology of the electronic system of intelligent and connected vehicle is becoming more and more complex, generally including multiple bus systems, and different bus systems are connected through gateways. On the one hand, gateway plays the role of bridging, converting two different bus protocols to each other; on the other hand, it also plays the role of gateway. It needs to control the data flow of the two bus systems according to practical application. Therefore, once the gateway programming logic vulnerabilities are found, cross-bus access attacks may be implemented to form the control of the vehicle system.
(4) There are more and more software such as on-board operating system/embedded system/on-board application, among which the proportion of products based on open source platform is increasing year by year. While it is beneficial for users to acquire more and more on-board software, it also buries the hidden danger of cyber security, which makes it easier for some IT personnel who do not have professional knowledge of automobile software to invade and control on-board system.

2.2 External Risks

External risk refers to the way an external attacker invades a car. Figure 2 shows the intrusive channel of typical intelligent network. These channels are all potential external security hazards of intelligent and connected vehicle, which can be used as a way to access the vehicle and launch malicious attacks. According to the different contact methods, it can be divided into two groups.

(1) Contact type, including OBD-II diagnostic port, mobile phone-locomotive connection parts, rear-mounted components, CD/DVD entertainment system, charging interface, etc.

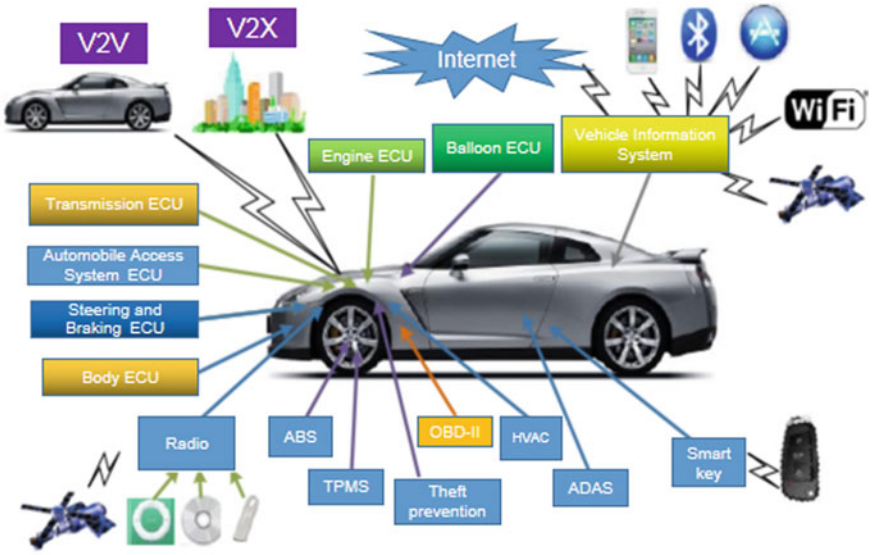


Fig. 2 Intrusive channels existing in the intelligent and connected vehicle

(2) Contactless, including sensor equipment (HD camera, Lidar, millimeter wave radar, GPS, etc.), short-range wireless (LTE-V/DSRC, TTMS tire pressure management system, wireless car key, vehicle Bluetooth, etc.), long-range wireless (Telematics, broadcasting, mobile App, remote update).

3 Analysis of Typical Attack Case

3.1 Example of Ford Winged Tiger Brake Attack

Automotive electronic components are connected by CAN network, and the communication between electronic components is carried out by CAN package. CAN is actually a large hub. When the CAN bus is idle, all units can start sending messages (multi-master control).

Figure 3 is the CAN bus network topology of the 2010 Ford Winged Tiger, whose CAN bus is more general, and has been attacked by hackers in real life, and representative. It can be seen that it has two CAN buses: a medium speed MS CAN (125 kbit/s) and a high speed HS CAN (500 kbit/s). Both buses are connected directly to the OBD-II diagnostic interface. That is, the vehicle can access the ECU directly from the OBD-II to the on-board bus. The high speed HS CAN bus is equipped with

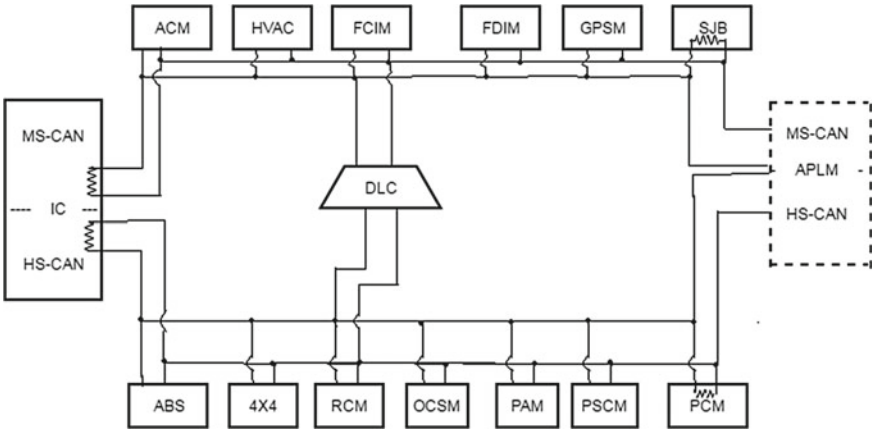


Fig. 3 Topology of 2010 Ford Winged Tiger CAN bus network

ABS anti-lock brake ECU and PSCM power. Steering control ECU, PCM power system control ECU and so on. ACM audio control ECU, HVAC HVAC ECU, FDIM pre-instrument time-limited ECU are installed on medium speed MS CAN bus [4].

(1) Physical access of vehicle CAN bus

The vehicle CAN bus is connected through the OBD-II maintenance diagnosis interface. Physical connection, the Ecom line is used to connect the computer through USB interface at one end and the automobile maintenance and diagnosis interface at the other end. When the connection is completed, the CAN bus can read and write data.

(2) CAN data acquisition

After the physical access is completed, the CAN bus data can be searched and captured by using EcomCat software. EcomCat software is written by two engineers in C language. It can detect and capture data in automobile CAN bus through OBD-II diagnostic interface, and provide CAN data filtering function. The purpose is to reduce the size of CAN data storage. The captured CAN data is saved as the "output.dat" file by default.

(3) CAN data analysis

CAN data "output.dat" file is captured from OBD-II maintenance diagnostic interface. By recording the sending time, data length and data content of Mframe data, the different actions of vehicle key ECU controller are analyzed, and the key CAN data frames are analyzed and decoded.

A set of data is intercepted from the output. dat file. According to the CAN bus message standard ISO 15765-2, the analytical research is shown in Fig. 4.

The first frame message: ID = 0x07E0, refers to the CAN message data sent to the ID OxO7EOECU unit, as shown in Fig. 5.

```
IDH: 07, IDL: E0, Len: 08, Data: 10 82 36 01 31 46 4D 43
IDH: 07, IDL: E8, Len: 08, Data: 30 00 00 00 00 00 00 00
IDH: 07, IDL: E0, Len: 08, Data: 21 55 30 45 37 38 41 4B
IDH: 07, IDL: E0, Len: 08, Data: 22 42 33 30 34 36 39 FF
IDH: 07, IDL: E0, Len: 08, Data: 23 FF FF FF 2A FF FF FF
```

Fig. 4 A set of data intercepted from the output.dat file

```
IDH: 07, IDL: E0, Len: 08, Data: 10 82 36 01 31 46 4D 43
IDH: 07, IDL: E8, Len: 08, Data: 30 00 00 00 00 00 00 00
IDH: 07, IDL: E0, Len: 08, Data: 21 55 30 45 37 38 41 4B
IDH: 07, IDL: E0, Len: 08, Data: 22 42 33 30 34 36 39 FF
IDH: 07, IDL: E0, Len: 08, Data: 23 FF FF FF 2A FF FF FF
```

Fig. 5 CAN message data is sent to the OxO7EOECU unit with ID

The first half byte "1" of the first data byte of the first frame message indicates that the message sent to the ECU unit with ID of 0x07E0 is multi-packet data. The second half byte of the first data byte and the second byte of the data constitute 0x082, which means that the payload data of the sending message is 0x082 bytes, and the third to eighth bytes are valid data.

The third frame to the fifth frame message is the continuous frame of the first frame (represented by the first half byte "2" of its data first byte). The second half byte of the first byte of the three consecutive frames data is "1", "2", "3". Therefore, after parsing, the actual data sent to the ECU with ID = 0x07E0 on the CAN bus is 3601 3146 D 4355 3545 3441 B 4233 3436 FF FH FF. FF 2AFF FF.

By analyzing the messages of Ford Winged Tiger under different operating conditions, the two engineers retrospectively analyzed the message control data of several key ECU units. Following is a brief list of Ford Wing Tiger brake instructions.

Based on the FORDISO14230.dll file (as shown in Fig. 6), the two engineers deduced and analyzed the CAN test data message, and excavated the key control instructions of Ford Winged Tiger brake (Fig. 7).

```
ISO14230_FORD_RequestDiagnosticDataPackets_A0          1000A730  52
ISO14230_FORD_DynanicallyDefineDiagnostikDataPacket_A1 1000A7A0  53
ISO14230_FORD_DiagnosticCommand_81                     1000A800  54
ISO14230_FORD_RequestCommonIdScalingMasking_24         1000A5C0  55
ISO14230_FORD_CommunicationControl_28                  1000A620  56
ISO14230_FORD_ControlDTCSetting_85                     1000A6E0  57
ISO14230_FORD_ReadMemoryByAddress_23                   1000A560  58
ISO14230_FORD_WriteMemoryByAddress_3D                  1000A680  59
```

Fig. 6 FORDISO14230.dll file

```
IDH:07, IDL:60, Len:08, Data;04 B1 00 3C FF 00 00 00
```

Fig. 7 Key control instruction for Ford Winged Tiger brake

The third and fourth bytes 0x003C of the data take part in the braking of Ford Wing Tiger, and a byte parameter, such as OxFF, is carried behind, indicating that the braking coefficient with OxFF parameter is applied.

(4) CAN data injection

By acquiring and parsing the CAN data message, the next step is to compile and inject the message into the automobile CAN bus by reverse engineering, in order to achieve the attack effect. Firstly, we need to define a set of CAN data message formats with effective behavior. The key code is shown in Fig. 8.

Initialization source code: handle = mydll. open_device (l, 0). Among them, "1" represents the high-speed CAN network, using 0 to select the first Ecom configuration line. Next, you can send CAN packets, as shown in Fig. 9.

The above code means that the "ID = 0x0230" data message is injected into the automobile CAN bus at a frequency of 1000 ms per time. In particular, when editing is completed, termination cyber should be included: mydll. close_device (handle).

From the above analysis, it is known that the brake control instructions of Ford Winged Tiger automobile are affected by injection of the instructions shown in Fig. 10.

```
mydll = CDLL('Debug\\ecomcat_api')
class SFFMessage(Structure):
    _fields_ = [("IDH",c_ubyte)],
    ("IDL",c_ubyte),
    ("data",c_ubyte*8),
    ("options",c_ubyte),
    ("DataLength",c_ubyte),
    ("TimeStamp,c_unit),
    ("baud",c_ubyte)]
```

Fig. 8 Key code of CAN data message format

```
y = pointer(SFFMessage())
mydll.DbgLineToSFF("IDH:02,IDL:30,Len:08,Data:A1 00 00 00 00 00 5D 30",y)
mydll.PrintSFF(y,0)
mydll.write_message_cont(handle,y,1000)
```

Fig. 9 Sends CAN packets

```
if do_diagnostic_session(mydll, handle, 0x760, "adj"):
    print "Started diagnostic session"

while True:
    print do_diagnostic_command(mydll, handle, 0x760, 0x3c, [0x7f])
```

Fig. 10 Attack directives

Once this message is injected into the CAN bus of the car, the Ford Winged Tiger is like a fixed car, even if you use how much energy to step on the accelerator pedal.

3.2 Typical Attacking Path Map of Intelligent and Connected Vehicle

With the development of intelligent and connected vehicle, attackers can attack more routes. In addition to close-range attack on vehicle brake system, attackers can also make long-range attack. Such attacks can take many forms, such as the attacker can completely suppress the effective braking, or the attacker can delay the braking or reduce the braking quality of the brake [3, 5]. The specific attack path is shown in Fig. 11.

According to the above description, if the disgruntled 4S shop repairman chooses the short-range attack mode, when the intelligent and connected vehicle upgrades through the cloud platform, it modifies the upgraded firmware and injects malicious programs to attack the ECU of the brake system; if it means to prove its technical strength, the hacker may choose the long-range attack mode and attack through DDoS, mainly to attack DSRC or 3G. In order to attack the ECU of the brake system ultimately, 4G is adopted. The attack tree is shown in Fig. 12.

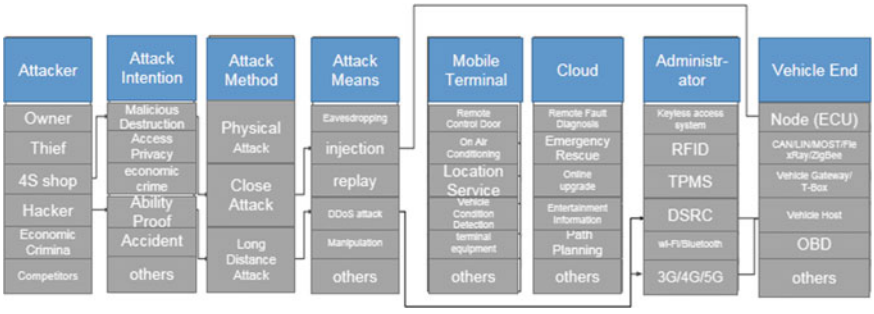


Fig. 11 The map of attacking brake function path

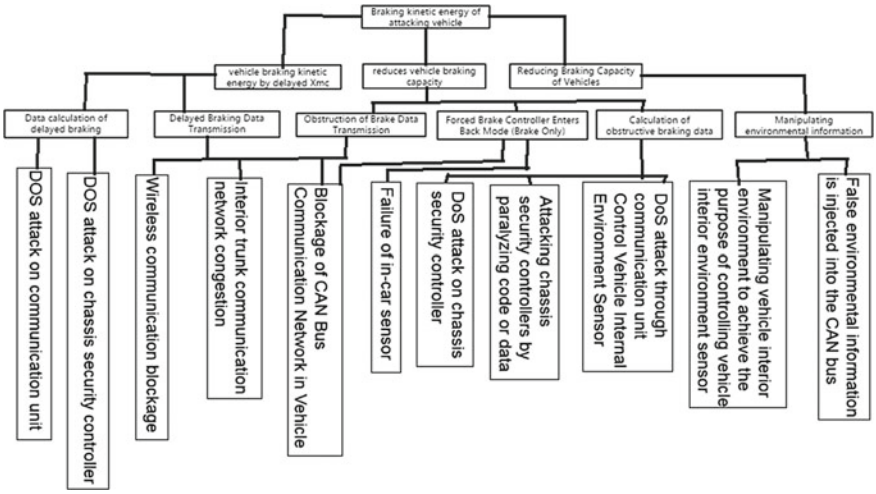


Fig. 12 Brake function attack tree

4 Conclusions

(1) The model of cyber security threat of an intelligent and connected vehicle is constructed, including internal risk and external risk. Internal risk mainly comes from ECU/MCU, bus, gateway and vehicular operating system/embedded system/vehicular application, etc. External risk refers to the way that external attackers invade the vehicle, which is divided into contact and non-contact.
(2) Based on the typical case of Ford Winged Tiger's brake system being attacked, the specific process of the brake system being attacked is analyzed by constructing the CAN bus network topology diagram of Ford Winged Tiger, according to the steps of CAN bus physical access, CAN data acquisition, CAN data parsing and CAN data injection.
(3) Typical attack path and specific attack tree of intelligent and connected vehicle are constructed.

Acknowledgements This research was supported by the Science and Technology Plan Project of Tianjin, China (17ZXRGGX00070, 18ZLZDZF00390).

References

1. Miller C, Valasek C (2013) Adventures in automotive networks and control units. Def Con 21:260–264
2. Miller C, Valasek C (2014) A survey of remote automotive attack surfaces. Black Hat USA

3. Miller C, Valasek C (2015) Remote exploitation of an unaltered passenger vehicle. Black Hat USA
4. Rutkowski L, Jaworski M, Pietruczuk L et al (2014) The CART decision tree for mining data streams. Inf Sci 266:1–15
5. Koscher K, Czeskis A, Roesner F, et al (2010) Experimental security analysis of a modern automobile. In: 2010 IEEE Symposium on Security and Privacy (SP). IEEE, pp 447–462

A Prediction Precision Inference Method for Passenger Alighting Station Based on the Condition Hypothesis

Fan Li, Qingquan Li, Zhao Huang and Jizhe Xia

Abstract Smart IC-card has been widely used in fare payment systems of public transport, which produces a large number of ticket checking records and spatiotemporal trajectory information. Accurately predicting passengers' travel stations based on IC-card data plays an important role in intelligent transportation. However, incomplete IC-Card transaction records are widely existing. The IC-card not only does not record the actual boarding stations but also lacks the information of alighting stations because passengers do not need to swipe card when they get off. Therefore, it is difficult to construct the actual passenger travel link, which makes it challenging to predict alighting stations accurately. Targeting on this challenge, we propose a "Boarding Cluster to Alighting Station" alighting station prediction model (BCTAS) by condition hypothesis. First, the model analyzes the travel characteristics of passengers' public transport. Second, the smart IC-card transaction records and map-matching algorithm are used to construct the mixed boarding station link. Third, the model performs the station clustering and cluster expansion to merge the same name station and the nearest station into a cluster, and further constructs the mixed boarding cluster link. Fourth, a Variable Order Markov Model that named Prediction by Partial Match (PPM) is adopted to predict the mixed boarding cluster link and then predict the boarding station. Fifth, the model infers the prediction precision of the alighting cluster and alighting station based on the condition hypothesis. Finally, our approach was evaluated by using the public transport data obtained in Shenzhen city, China. The results show that (a) with the increase of training data, the precision of the model is gradually enhanced, (b) by using the mixed boarding cluster link, the prediction precision of the boarding cluster and boarding station could reach 88.05%

F. Li · Q. Li · Z. Huang · J. Xia (✉)
Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China
e-mail: 331650353@qq.com

F. Li
College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China

Z. Huang
College of Information Engineering, Shenzhen University, Shenzhen 518060, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_81

and 84.52% respectively, (c) Based on the condition hypothesis, it can be inferred that the lower limit of the prediction precision of the alighting cluster and alighting station is 78.09% and 74.96%, respectively.

Keywords Alighting station prediction · Smart IC-card transaction records · Station clustering and cluster expansion · Variable order Markov model · Prediction by partial match model (PPM) · Condition hypothesis

1 Introduction

With the development of the intelligent public transportation system, smart IC-card is widely used in fare payment system of public transport, resulting in a large number of transaction records. Although the mainstream travel survey is the basis for establishing the traffic demand model and developing traffic planning, it is a very resource-consuming work with large errors and unsatisfactory results. Based on smart IC-card transaction records to predict passenger boarding and alighting stations is one of the most potential methods to predict public transport distribution, which is an important work for data cleaning and reuse. The distribution prediction of bus travel can greatly reduce the financial expenditure of government traffic survey. In the commercial field, the prediction also plays an important role such as the game of shared bicycle, accurate bus travel prediction and reasonable bicycle delivery can maximize the profits of related enterprises. Finally, in the aspect of urban planning and management, accurately predict the passenger boarding and alighting stations is not only helpful to master the change of bus passenger flow, assist the decision of bus operation and optimize bus line, but also provides a new perspective for the study of individual/group public transport travel characteristics, the estimate of urban traffic flow, urban spatial structure and functional adaptability analysis [1–3].

In recent years, due to the development of communication technology, a large number of high-precision personal travel trajectory data based on the intelligent transportation system, mobile communication, and social network has been collected. These datasets were widely used to understand better different human travel behavior such as travel distance distribution feature, moving track radius of rotation, and probability density distribution of the access area [4–8]. Based on the recent understanding of human mobility patterns, researches have been done to study the predictability of human motion patterns [6, 9, 10]. Some researches show that human motion could be highly regular in certain spatial scales [6], so human behaviors can be accurately recognized [11] and predicted [12]. Based on this, some studies have been conducted to predict personal next location based on users' trajectories [13]. Generally, location predictions can be classified into the continuous trajectory [14, 15] and the discontinuous trajectory [16], which includes the aggregate model and the individual selection model. The specific methods include the model based on the historical trajectory, the Markov model, the Bayesian inference model, and the Integrated model [17–20]. Articles show that the Markov Model can effectively solve

the problem of location prediction [21–25]. The traditional N-order Markov (NM) Model [21, 22] and the Hidden Markov Model [23] use the state transition matrix to calculate the transition probability and then predict. However, the relatively high spatial complexity and zero frequency because of sparse data limit their development. To solve these two problems, the Variable Order Markov Model is proposed such as the Prediction by Partial Match (PPM) algorithm [24, 25].

However, the IC-card transaction records are often difficult to be used by the relevant departments, resulting in a large waste of data. The reason is (1) the smart IC-card usually only records information such as passenger's ID, transaction time, bus plate number and other information. The passenger's actual boarding stations have not been explicitly recorded. (2) In addition to subway records, a large number of bus lines in many large and medium-sized cities in our country adopt the mode of swiping card once, so passengers do not need to swipe when they get off, which is considered as an "incomplete trading mode" by the industry. Therefore, the construction of passenger travel link is tough, making it an extremely challenging task to predict alighting stations accurately. Overall, the academic study on the prediction of passenger alighting stations based on smart IC-Card is still in the exploratory stage [26–30], and the related researches are not deep enough. Targeting on this issue, we propose a novel approach to predict passenger alighting stations by condition hypothesis. This paper is organized as follows: Sect. 2 provides a view of the data used in this research. Section 3 outlines the mechanism of "Boarding Cluster to Alighting Station" alighting station prediction model (BCTAS). Section 4 presents the experimental results and discusses existing issues. Section 5 provides conclusions.

2 Data Description

The dataset used in this research were obtained in Shenzhen city, China, covering a total of 30 days in September 2014. There are three kinds of data used in this dataset: Bus station and line vector data, bus trajectory data and smart IC-card transaction records (summarized in Table 1):

(1) Station data and line data: The data is obtained from the Gaode Map through Open API. The subway data includes 118 stations on five lines. The station data contains station number, station latitude and longitude, station name and line of the station. The bus line data collects the detailed geographic information of 521 bus lines in Shenzhen City, including 270 up lines and 251 down lines. Among the 16,082 bus stations on 521 bus lines, there are 6817 pairs of same name stations in the up and down lines and 3155 stations after merging the same name stations.

(2) Bus trajectory data: The data provided by Shenzhen's Comprehensive Traffic Operation Command Center was collected on the GPS terminal device that installed on 14,484 buses in Shenzhen City, including bus plate number, time

Table 1 A summary of the experiment

Data type	Description	Features	Data source
Station data and line data	Specific information about 16,082 bus stations, 521 bus lines, 118 subway stations and 5 subway lines in Shenzhen	Station number, station latitude and longitude, station name, the line of the station, 270 up lines, 251 down lines	Obtained from the Gaode Map through Open API
Bus trajectory data	The trajectory of 14,484 buses in Shenzhen	Bus plate number, time of data acquisition, instantaneous latitude and longitude, instantaneous speed, instantaneous direction angle	Provided by Shenzhen's Comprehensive Traffic Operation Command Center
Smart IC-card transaction records	Bus and subway transaction records	Passenger's anonymous ID, time of swiping card, travel mode, subway station or bus plate number, consumer price	Provided by Shenzhen Tong company

of data acquisition, instantaneous latitude and longitude, instantaneous speed, and instantaneous direction angle. Most of the bus trajectories are sampled at low frequency with a sampling period of about 60 s.

(3) Smart IC-card transaction records: The data provided by Shenzhen Tong Company is anonymous and does not record personal information. The data consists of bus and subway transaction records, including the passenger's anonymous ID, time of swiping card, travel mode, subway station or bus plate number, and consumer price. There are 4,524,174 passengers chose public transportation with IC-card this month, and the number of swiping card is 56,247,778.

3 Passenger Alighting Station Prediction

The "Boarding Cluster to Alighting Station" alighting station prediction model (BCTAS) predicts passenger alighting stations by condition hypothesis with the following processes (Fig. 1): (1) analyzes the travel characteristics of passengers' public transport, (2) the smart IC-card transaction records and map-matching algorithm are utilized to construct the mixed boarding station link, (3) clusters the bus and subway stations and expands the clusters to merge the same name station and the nearest station into a cluster, and the mixed boarding station link is transformed into the mixed

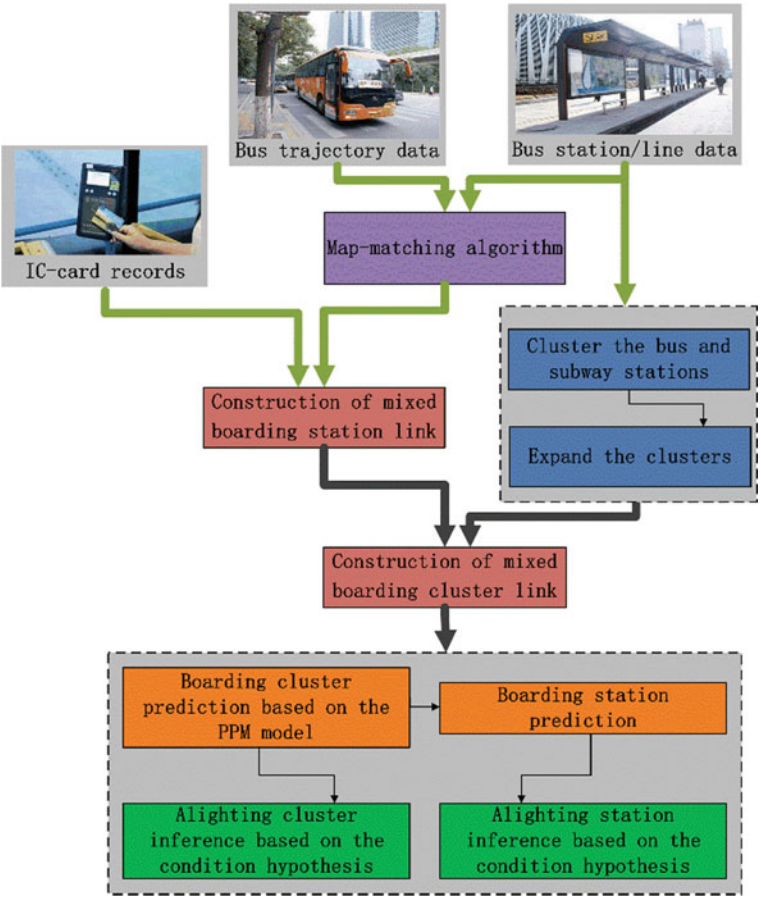


Fig. 1 "Boarding Cluster to Alighting Station" alighting station prediction model

boarding cluster link, (4) a Variable Order Markov Model which named Prediction by Partial Match (PPM) is adopted to predict the mixed boarding cluster link, and further predict the boarding station, and finally (5) infers the prediction precision of the alighting cluster and alighting station based on the condition hypothesis.

3.1 Travel Characteristics of Public Transport

Public transport mainly includes subway and bus. A complete travel link can be a single travel mode or a combination of multiple travel modes. Figure 2 shows the proportion of passengers with different bus travel proportions and the number of passengers with different daily average transfer times. The results show that the

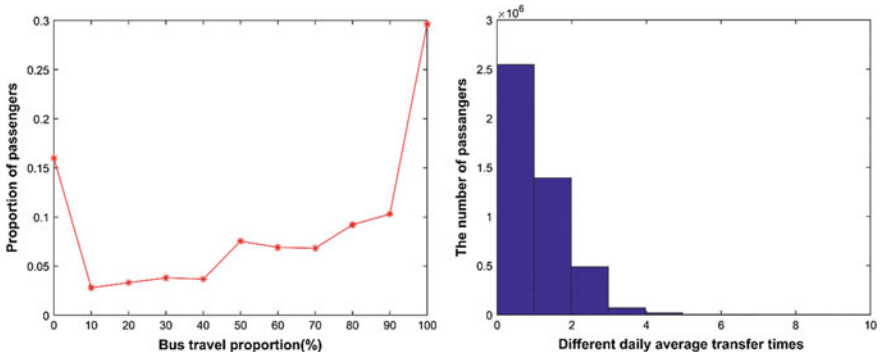


Fig. 2 Travel characteristics of public transport

proportion of passengers who only chose to travel by bus is 29.62 and 54.41% of the passengers chose two travel modes in a month, which indicates that more than half of the passengers adopt the travel mode of "bus + subway". It can be seen that 28.76% of the passengers have a bus travel proportion of 30–70%, which usually switch or transfer between two travel modes. The statistical results of transfer times show that each passenger transfers an average of 0.75 times a day. Therefore, the mixed travel link combining bus data and subway data is complete and can better reflect the travel behavior of passengers.

3.2 Construction of Mixed Travel Link

3.2.1 Construction of Mixed Boarding Station Link

The construction of mixed boarding station link requires boarding station data of subway and bus. Since IC-cards do not record the complete boarding stations, we need to match the instantaneous location of swiping card to the bus station according to the bus trajectory and the location information of bus stations. The first step is map matching, which matches the location of swiping card to the bus trajectory. Since the original bus trajectory data is sampled with low frequency, in this paper, we adopt a map-matching approach for low frequency floating car, which is called MDP-MM to restore the bus trajectory [31]. Then, based on the average speed interpolation algorithm, the bus trajectory is interpolated uniformly by second, and the instantaneous locations of swiping card are matched to the bus trajectory according to the time of swiping card. The second step is station matching, which use the nearest neighbor rule, the most common way in bus station matching, to take the bus station that nearest to the location of swiping card as the boarding station based on the location information of bus stations. Finally, the bus and subway boarding stations of each

passenger are connected to a mixed boarding station link [S1 → S2 → ··· → Sn] in chronological order.

3.2.2 Construction of Mixed Boarding Cluster Link

When a passenger gets off at station A, there are usually two cases for the passenger to get on the next station: (1) changes the original direction and select the same name station A* of station A or the nearest station B to return; (2) travels along the original direction and select the current station A forward. Due to the sparseness of transaction records and the diversity of station selection, to increase the stability of transition probability between boarding stations, this paper spatially clusters the bus and subway stations to merge all the stations in the adjacent space domain into a meaningful location. The principle of clustering is not only to ensure that the same name stations and the nearest station of the up and down lines are clustered together but also to effectively divide the adjacent stations of the same line into different clusters. K-Means [32] and DBSCAN [33] are two of the most widely used algorithms for clustering. Due to the nature of station data, we chose the DBSCAN approach to reduce the impact of unbalanced data distribution and detect non-convex clusters. Therefore, a DBSCAN clustering process is used to assign those stations around the core points that satisfying the density threshold to the same cluster according to the density connectivity between stations. After clustering, since it is not guaranteed that for all bus stations A, the nearest station B and the same name station A* are already in the same cluster with A, so a cluster expansion is needed. The expansion traverses all bus stations to see if A* and B are already in the same cluster with A. If not, the cluster of A* and B will be merged with the cluster of A. After clustering and expansion, the mixed boarding station link [S1 → S2 → ··· → Sn] is transformed into the mixed boarding cluster link [C1 → C2 → ··· → Cm].

3.3 Prediction by Partial Match Model

The PPM algorithm uses the "Dictionary Tree" to solve the problem of low space utilization, and utilizes the escape mechanism to solve the zero frequency problem. To build a dictionary tree of clustering link, the order D of the model should be determined. For station prediction problem, if a cluster does not appear after the training clustering link, the escape mechanism is used. For each clustering link s of length k($k \leq D$), the escape mechanism assigns a conditional probability $\hat{P}(escape|s)$ to all clusters which did not appear after the clustering link s in the training sequence, and the remaining probability $1-\hat{P}(escape|s)$ is distributed in other clusters which appear after the clustering link s. The following equation shows the mechanism:

$$P(c|s) = \begin{cases} \hat{P}(c|s), & if\ c \in \Sigma s \\ \hat{P}(escape|s) \cdot P\left(c|s'\right), & otherwise \end{cases} \tag{1}$$

For the empty clustering link ε, PPM takes

$$P(c|\varepsilon) = \frac{1}{|\Sigma|} \tag{2}$$

where $\sum$ denotes all clusters that appear in the training set, and c indicates any cluster in the cluster set. The $\sum$s in the equation denotes all clusters which appear after the clustering link s in the training set, and s' denotes the longest suffix of the clustering link s.

For each clustering link s and cluster c, N(sc) represents the number of occurrences of sc in the training set. Let $\sum$s be the cluster set of the training set that appears after the clustering link s, that is, $\sum s = \{c: N(sc) > 0\}$. The following are related equations:

$$\hat{P}(c|s) = \frac{N(sc)}{\left|\sum s\right| + \sum_{c' \in \Sigma s} N(sc')}, if\ c \in \sum s \tag{3}$$

$$\hat{P}(escape|s) = \frac{|\Sigma s|}{|\Sigma s| + \sum_{c' \in \Sigma s} N(sc')} \tag{4}$$

In the learning phase, the PPM algorithm builds a dictionary tree T from the training set. Each node in T is related to a cluster and has a counter. If it is determined the order is D, then the maximal depth of T is D + 1. The algorithm starts from the root node corresponding to the empty cluster link Ɛ and gradually parses the training sequence to one cluster c and its sub-link of size D at a time, which usually define a path in tree T. After parsing the first D clusters, the length of each newly constructed path is D + 1. The counters along this path are incremented. Therefore, the counter of any node denotes the number of occurrences that cluster c (corresponding to the node) appears after the clustering link s represented by the path, that is N(sc). After building the dictionary tree T, we can calculate the conditional probability P(c|s) by that dictionary tree, where c is any cluster in the cluster set and s is a clustering link whose length is less than or equal to D. In the calculation process, we first traverse the dictionary tree from the root node, and the traversal rule is whether s', the longest suffix of s matches. After that, s'c denotes the complete path from the root node to the leaf, and then (2)–(4) can be used to calculate the conditional probability.

When predicting the next boarding cluster, for each passenger, all IC-card transaction records can form a mixed boarding station link. After spatial clustering and expansion, the mixed boarding station link is transformed into a mixed boarding cluster link. We use each passenger's cluster link to train the PPM model and calculate the probability of passenger's appearance in each possible spatial domain in the future based on the current cluster. Finally, we make the spatial domain of maximum probability as the prediction result of the boarding cluster.

3.4 Condition Hypothesis

Due to the lack of concrete alighting stations from smart IC-card, the prediction of alighting stations could be transformed into the prediction of boarding stations. Generally, when passengers get off at some station, they tend to choose a nearby station for the next boarding whether they change the original direction or move in the same direction. Therefore, the following condition hypothesis is made in the design of the prediction model: the next boarding station chosen by the same passenger is usually the current alighting station, the same name station of the alighting station or other stations in the adjacent space domain in the opposite direction. Since not all bus stations have same name station, and among the 6817 pairs of same name stations, there are 372 stations, their nearest stations in the opposite direction are not their same name stations, other neighbor stations need to be considered. However, there are some errors in this hypothesis, so this paper uses the subway data to evaluate the errors in the experiment.

4 Experiment and Results

4.1 Cluster Stations and Expand Clusters

For the DBSCAN algorithm, the MinPts parameter of the minimum numbers of points is set to two. Figure 3 shows that when the MinPts is set to two and the Eps parameter of the cluster radius increases from 100 to 500 m, the proportion of the same name stations are already in the same cluster increases from 82.89 to 99.95%. When the cluster radius is 150 m, 91.42% of the same name stations are already in the same cluster, and the cluster radius is far less than the average station distance of 654.66 m in Shenzhen city, which can effectively distinguish the adjacent stations of the same line from different clusters, so the cluster radius is set to 150 m. To merge each nearest station and the same name station of the up and down lines into the same cluster, we need to expand the clusters spatially. After the first step of station clustering, all bus stations are clustered into 2104 clusters. After the second step of cluster expansion, there are 1554 clusters left. Figure 4 shows the clustering result after cluster expansion.

4.2 The Prediction of the Boarding Cluster and Boarding Station

Location prediction technology requires a large amount of data to train the model. The more data used for training, the more often the model can reflect the passenger travel rule. Since people travel more regularly on weekdays than on weekends, our

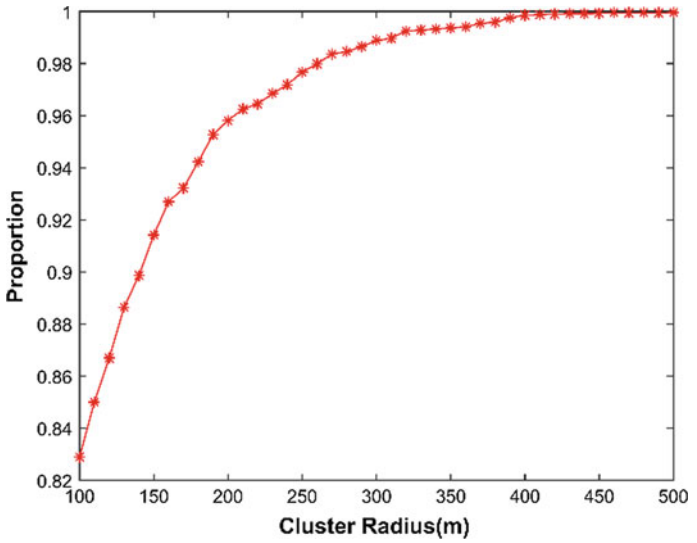


Fig. 3 The relation between clustering result of the same name stations and the cluster radius

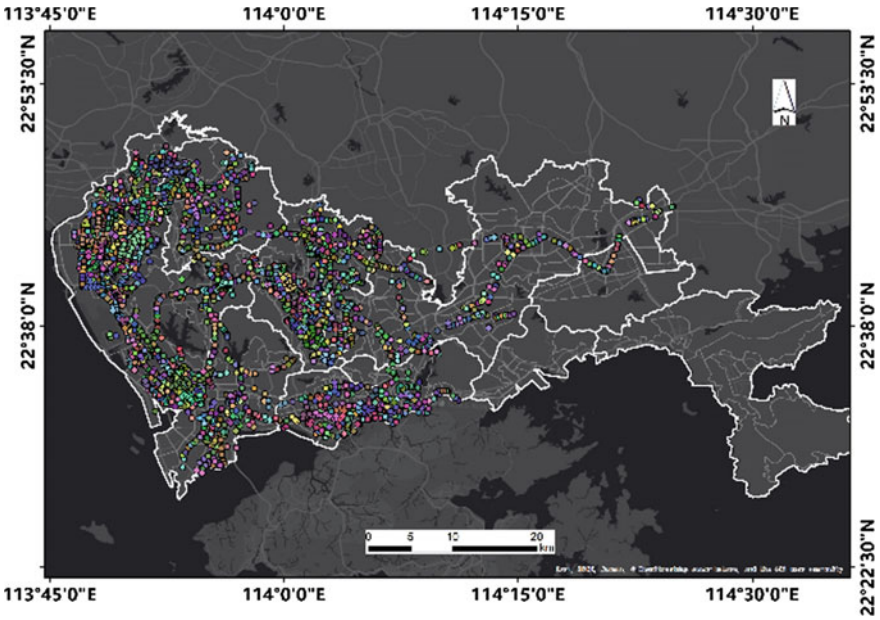


Fig. 4 The clustering result after cluster expansion

experiment first chose people who swiped cards for 5–21 weekdays in September 2014. Then the IC-card transaction records of each passenger are converted into the mixed boarding station link and further converted into the mixed boarding cluster link. Finally, the PPM models with different orders are used to train and calculate the cluster prediction precision of the cluster model when the sample data increases from 5 weekdays to 21 weekdays, in which different orders represent the number of known continuous boarding clusters. Figure 5 shows that the cluster prediction precisions of two cluster models are improved with the number of workdays gradually increasing. It can be seen that the cluster prediction precision of the one-order PPM boarding cluster model increases from 54.49 to 85.42%, and the precision of the two-order boarding cluster model increases from 75.58 to 88.05% from 5 workdays to 21 workdays. For Variable Order Markov Model, since they can adaptively adjust according to the length of the input sequence to find the sub-sequence that best matches it, the prediction precision of the two-order PPM cluster model is higher than that of the one-order PPM cluster model.

For the prediction of the boarding station, the station with the highest probability in the predicted cluster is chosen as the predicted boarding station in the experiment. Figure 6 shows that from 5 workdays to 21 workdays, the station prediction precision of the one-order boarding station model increases from 45.46 to 82.77%, and the precision of the two-order boarding station model increases from 62.11 to 84.52%.

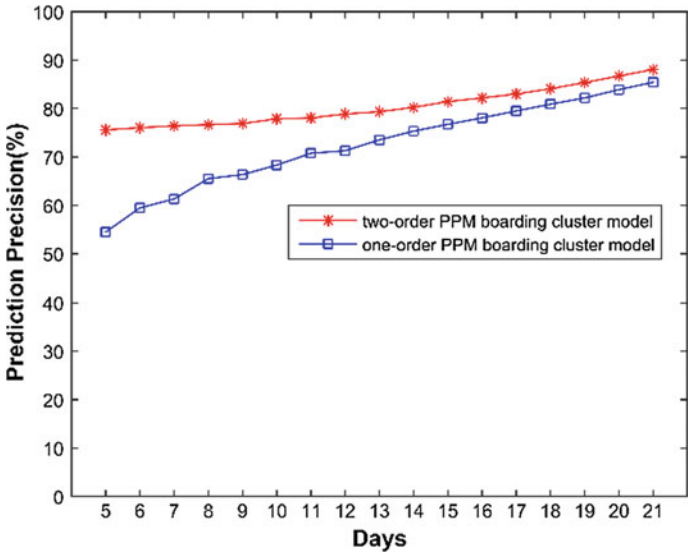


Fig. 5 Prediction precision of the boarding cluster model under different orders

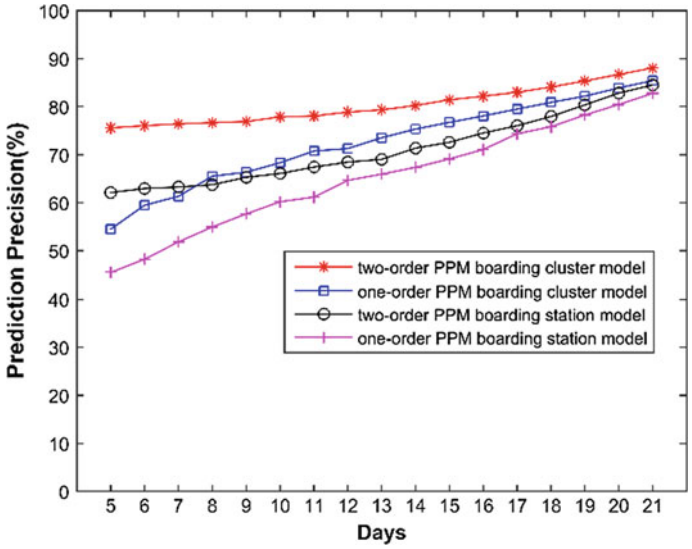


Fig. 6 Prediction precision of the boarding cluster/station model under different orders

4.3 The Inference of the Alighting Cluster and Alighting Station

The previous article assumes that the next boarding station chosen by the same passenger is usually the current alighting station, the same name station of the alighting station or other stations in the adjacent space domain in the opposite direction. In order to verify the hypothesis, this paper chose passengers who have subway records in 21 workdays to construct the mixed boarding and alighting link and count the repeatability of boarding stations and alighting stations. The result shows that the repeatability of the predicted alighting cluster and actual alighting cluster reaches 88.69% if the next boarding cluster is used as the current alighting cluster, which indicates that the next boarding station and the current alighting station are usually in the same cluster. Since the clustering and expansion steps have clustered the alighting station, the same name station, and the nearest station into the same cluster, the next boarding station is also in the same cluster with the current alighting station's same name station and nearest station. According to the principle of clustering, we can know that the next boarding station may be the current alighting station, or the same name station and other adjacent stations in the opposite direction, which verifies our hypothesis.

The verification of the condition hypothesis indicates that if the next boarding cluster is used as the current alighting cluster, it can be considered that the prediction precision of the alighting cluster is reduced by 11.83% compared with that of the boarding cluster. Therefore, using the data of 5–21 workdays, the prediction precision of the one-order PPM alighting cluster model can be inferred to be between 48.33 and

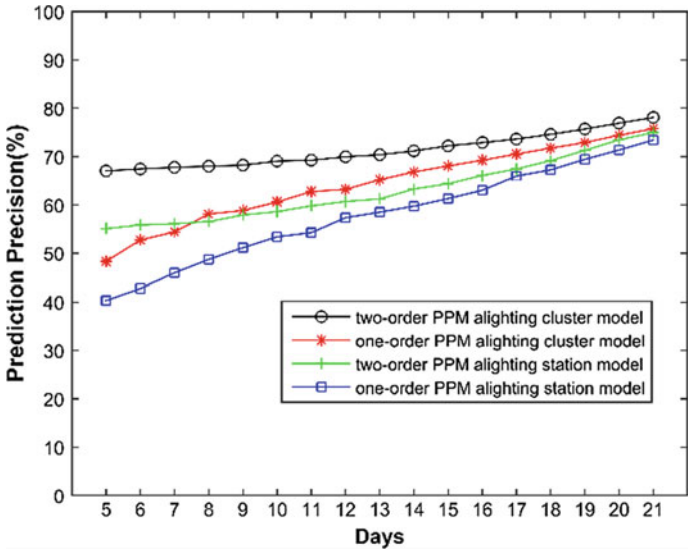


Fig. 7 Prediction precision of the alighting cluster/station model under different orders

75.76%, and the precision of the two-order alighting cluster model can be inferred to be between 67.03 and 78.09%. For the prediction of the alighting station, based on the condition hypothesis, it can be inferred that the prediction precision of the one-order PPM alighting station model is between 40.32 and 73.41%, and the precision of the two-order alighting station model is between 55.09 and 74.96%. The results are shown in Fig. 7. Compared with the prediction of the boarding station, the prediction of the alighting station has a constraint about the boarding line. When predicting the alighting station, the station with the highest probability on the boarding line in the next boarding cluster can be used as the predicted alighting station. If the next boarding cluster does not include any station on the current boarding line, then the station with the highest probability except the boarding station on the current line is taken as the predicted alighting station.

4.4 Discussion

As shown in Fig. 8, the prediction precision of the boarding station is lower than that of the boarding cluster. This is because our experiment only selects the station with the highest probability in the predicted cluster based on the prior probability without any other condition, thus reducing the prediction precision. On the other hand, the prediction precisions of the alighting cluster and alighting station are only inferred values because it is difficult to obtain the real boarding stations of passengers

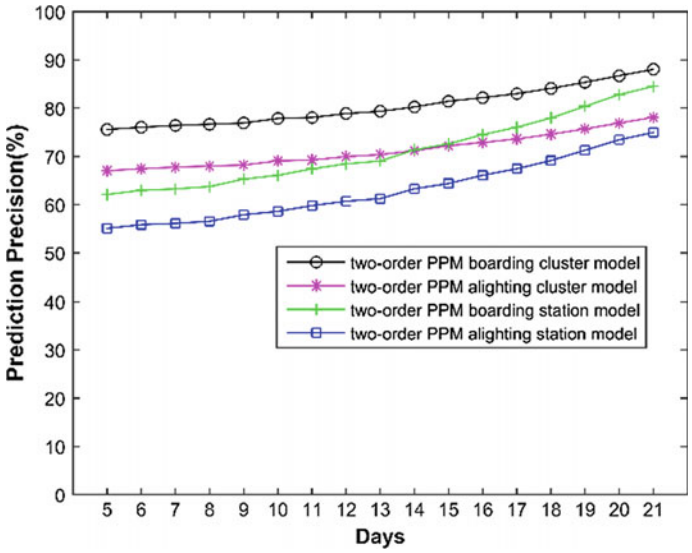


Fig. 8 Prediction precision of the boarding cluster/alighting cluster/boarding station/alighting station model

on a large scale in practice. Therefore, the prediction of alighting stations is transformed into the prediction of boarding stations in our experiment, and the condition hypothesis about the next boarding stations is made to support this conversion. In the process of verifying the hypothesis, we have found the high cluster repeatability of the boarding and alighting station, which can be used to infer the prediction precision of the alighting cluster and alighting station. Finally, the prediction precision of the alighting station is lower than that of the boarding station. Generally, the prediction performance of the alighting station can also be as excellent as that of the boarding station if the next boarding station is the three kinds of stations mentioned in the condition hypothesis. However, the repeatability of the boarding cluster and the alighting cluster is 88.69%, which leads to the prediction performance of alighting station lower than that of the boarding station. Due to the line constraint of the alighting station, in fact, the theoretical prediction precision of the current alighting station is higher than the inferred value, that is, the prediction precision of the alighting station is a theoretical lower limit. In the future, we will try to find the real survey data of bus passengers' boarding and alighting stations to verify the results, to further evaluate the performance of BCTAS model.

5 Conclusion

This paper proposes a "Boarding Cluster to Alighting Station" alighting station prediction model (BCTAS) by condition hypothesis. First, BCTAS analyzes the travel characteristics of passengers' public transport. Second, the model utilizes the smart IC-card transaction records and map-matching algorithm to construct the mixed boarding station link. Third, the BCTAS clusters the bus and subway stations and expands the clusters to merge the same name stations and the nearest station into a cluster, and further construct the mixed boarding cluster link. Fourth, a Variable Order Markov Model that named Prediction by Partial Match (PPM) is adopted to predict the mixed boarding cluster link, and then the boarding station is predicted. Fifth, the model infers the prediction precision of the alighting cluster and alighting station based on the condition hypothesis. Finally, the experimental results demonstrate strong predictability in our study area, and the prediction precision of the BCTAS is gradually enhanced with the increase of training data. By utilizing the mixed boarding cluster link, the prediction precision of the two-order boarding cluster model and two-order boarding station model reach 88.05% and 84.52% respectively. Based on the condition hypothesis, it can be inferred that the lower limit of the prediction precision of the two-order alighting cluster model and two-order alighting station model is 78.09% and 74.96%, respectively.

References

1. Medina SAO, Erath A (2013) Estimating dynamic workplace capacities by means of public transport smart card data and household travel survey in Singapore. Transp Res Record J Transp Res Board 2344(-1):20–30
2. Long Y, Shen Z (2015) Finding public transportation community structure based on large-scale smart card records in Beijing. Geospatial Analysis to Support Urban Planning in Beijing. Springer International Publishing
3. Zhong C, Arisona, SM et al (2014) Detecting the dynamics of urban structure through spatial network analysis. Int J Geogr Inf Sci 28(11):2178–2199
4. Brockmann D, Hufnagel L, Geisel T (2006) The scaling laws of human travel. Nature 439(7075):462–465
5. Gonzalez MC, Hidalgo CA, Barabasi A-L (2008) Understanding individual human mobility patterns. Nature 453(7196):779–782
6. Song C, Qu Z, Blumm N et al (2010) Limits of predictability in human mobility. Science 327(5968):1018
7. Jiang B, Yin J, Zhao S (2009) Characterizing the human mobility pattern in a large street network. Phys Rev E: Stat, Nonlin, Soft Matter Phys 80(1):1711–1715
8. Roth C, Kang SM, Batty M et al (2011) Structure of urban movements: polycentric activity and entangled hierarchical flows. PLoS ONE 6(1):e15923
9. Lin M, Hsu WJ, Zhuo QL (2012) Predictability of individuals' mobility with high-resolution positioning data. In: ACM Conference on Ubiquitous Computing, pp 381–390
10. Lian D, Zhu Y, Xie X et al (2014) Analyzing location predictability on location-based social networks. Adv Knowl Discovery Data Mining, 102–113
11. Kuge N, Yamamura T, Shimoyama O et al (2000) A driver behavior recognition method based on a driver model framework. SAE Trans 109(6):469–476

12. Pentland A, Liu A (1999) Modeling and prediction of human behavior. Neural Comput 11(1):229–242
13. Zheng X, Han J, Sun A (2018) A survey of location prediction on Twitter. IEEE Trans Knowl Data Eng 30(9):1652–1671
14. Scellato S, Musolesi M, Mascolo C et al (2011) NextPlace: a spatio-temporal prediction framework for pervasive systems. In: International conference on pervasive computing. Springer, Berlin, pp 152–169
15. Du Y et al (2018) A geographical location prediction method based on continuous time series Markov model. PLOS ONE 13(11)
16. Noulas A, Scellato S, Lathia N et al (2012) Mining user mobility features for next place prediction in location-based services. In: IEEE international conference on data mining. IEEE, New York, pp 1038–1043
17. Li Q, Zheng Y, Xie X et al (2008) Mining user similarity based on location history. In: ACM Sigspatial international conference on advances in geographic information systems. ACM, New York, p 34
18. Jeung H, Liu Q, Shen HT et al (2008) A hybrid prediction model for moving objects. In: Proceedings of the 24th IEEE international conference on data engineering. IEEE Press, Cancun, Mexico, pp 70–79
19. Do TMT, Gatica-Perez D (2012) Contextual conditional models for smartphone-based human mobility prediction. In: ACM conference on ubiquitous computing. ACM, New York, pp 163–172
20. Montoliu R, Blom J, Gatica-Perez D (2013) Discovering places of interest in everyday life from smartphone data. Multimedia Tools Appl 62(1):179–207
21. Ashbrook D, Starner T (2003) Using GPS to learn significant locations and predict movement across multiple users. Pers Ubiquit Comput 7(5):275–286
22. Gambs S, Killijian M-O et al (2012) Next place prediction using mobility Markov chains. In: EUROSYS 2012 workshop on measurement, privacy, and mobility, p 3
23. Mathew W, Raposo R, Martins B (2012) Predicting future locations with hidden Markov models. In: ACM conference on ubiquitous computing. ACM, New York, pp 911–918
24. Begleiter R, El-Yaniv R, Yona G (2011) On prediction using variable order Markov models. J Artif Intell Res 22(1):385–421
25. Yang J (2015) Research on location prediction based on historical trajectory. Hangzhou University of Electronic Science and Technology
26. Hu J, Deng J, Huang Z (2014) A judgment probability model of the alighting stations of the passengers with the bus IC card based on the trip link. Transp Syst Eng Inf 14(2):62–67
27. Li D, Lin Y, Zhao X et al (2011) Estimating a transit passenger trip origin-destination matrix using automatic fare collection system. In: Database systems for advanced applications. Springer, Berlin, Heidelberg, pp 502–513
28. Zhang F, Yuan NJ, Wang Y et al (2015) Reconstructing individual mobility from smart card transactions: a collaborative space alignment approach. Knowl Inf Syst 44(2):299–323
29. Jiayi L, Jin Z, Jingwen Z et al (2018) An algorithm to identify passengers' alighting stations and the effectiveness evaluation. Geomatics and Information Science of Wuhan University
30. Yilin W, Zhjgang J (2017) Individual station estimation from smart card transactions. J East China Normal Univ (Natural Science) 05:210–221
31. Chen BY, Yuan H, Li Q et al (2014) Map-matching algorithm for large-scale low-frequency floating car data. Int J Geogr Inf Sci 28(1):22–38
32. Macqueen J (1965) Some methods for classification and analysis of multivariate observations. In: Proceedings of Berkeley symposium on mathematical statistics and probability, pp 281–297
33. Ester M, Kriegel H P, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: International conference on knowledge discovery and data mining. AAAI Press, Palo Alto, pp 226–231

Deep-Learning-Based Detection of Obstacles in Transit on Trams

Yiming Li and Guoqiang Cai

Abstract Due to the pervasive employment of trams, the measurements to keep the transit of trams safe are necessary and emergent. In this sense, we put forward a Neural Network based on Convolutional Neural Network, in which we made some modifications to make it more flexible. Such as passthrough layer to ensure some small objects detectable, anchor boxes to ensure a high-speed detection, and batch-normalization layers to make the network be malleable for objects with different distributions. With this network, we can efficiently detect possible obstacles, such as pedestrians, cars and some other objections that may endanger the trams. We test the network among several databases with 5000 samples, and the average accuracy rate is 94.12%, the average detecting speed is 30 FPS, the smallest detectable object's size is 20 × 20 pixels, these all show remarkable result.

Keywords Deep learning · Tram · Real-time detection · Batch normalization

1 Introduction

With the proceed of urbanization, as for solving seriously urban traffic problems, trams with safe, eco-friendly and cheap characters are gradually taken into public's account. But with the gradually more increased trams are used to release transport pressures, the safety of the trams while operating renders more public's concerns over as well.

Nowadays trams are continuously becoming one of the most important improved points for intelligent transportation. Compared with urban subways, the environment for trams to operate is more independent, but they cannot take the road personally [1]. In this case, the uncivil or illegal behaviors will take place occasionally, which will lead to some obstacles unpredictable. Thus, making efforts to find a way to detect the

Y. Li · G. Cai (✉)
Beijing Jiaotong University, Beijing 100044, China
e-mail: gqcai@bjtu.edu.cn

State Key Lab of Rail Traffic Control and Safety, Beijing 100044, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_82

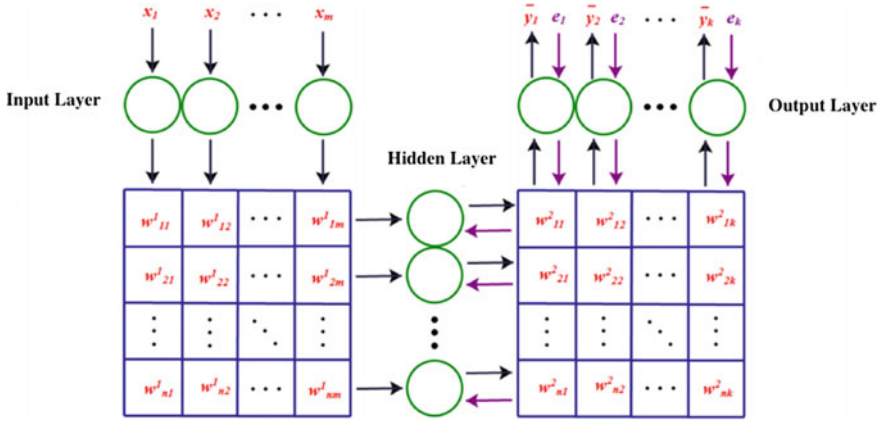


Fig. 1 The illustration about the structure of neural network

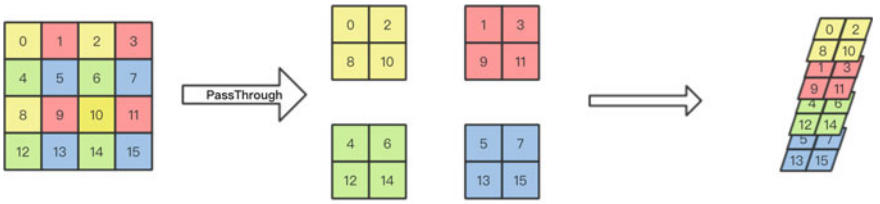
front obstacles timely is meaningful and practical for the safety of trams' operations (Fig. 1).

In recent years, the way to detect objections by Convolutional Neural Network (CNN) is on the hot point. Based on its fast detecting speed and significant accuracy, the CNN has been used in many fields, such as face detection, X-ray detection and behaviors detection [2]. Our research bases on the CNN's original structure and, in order to get great performance, we make some modifications and additions to the network to make sure that the detection speed and accuracy can satisfy the requirements.

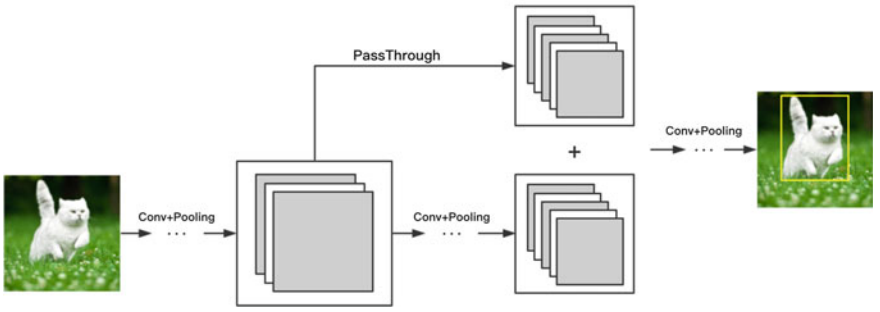
2 Passthrough Layer for Tiny Obstacles

Due to the high speed of trams and high detection accuracy that required, the network needs to be able to detect tiny or far obstacles possibly and timely. The normal CNN has its limits that cannot effectively extract the small or insignificant characters, which stand for the possibility to detect and figure out small objections in a picture. In this case, we use a special layer called 'Passthrough Layer' to ensure that the network has abilities to extract more detailed characters of different obstacles. Thus, the networks can get more details of the tiny obstacles.

Normally, the CNN extracts and compresses image features by compressing the feature map size on the two-dimension (length and width) and deepening the depth of the channel to reduce the detection pressure of the object detection and positioning layer at the end part of network [3, 4]. Obviously, this method can promote the efficiency of the network, but due to the compression of characters in two-dimension, most of the characters are represented by channel depth. Though they are enough to detect big-size objections in most conditions, however, for objects with relatively



(a). The way how PassThrough Layer works



(b). The exact position of PassThrough Layer in a network

Fig. 2 The realization of PassThrough layer

small scales in the original image, they are easily ignored in the case of insufficient depth since the features on the two-dimensional level are expressed in the form of channel depth. In this condition, the recognition result is often not ideal.

The essence of the PassThrough layer is feature's rearrangement. For instance in Fig. 2a, the feature map of 4 * 4 * 1 are sampled by the row and column respectively, and four feature maps of 2 * 2 * 1 are obtained, which are connected in series according to the channel to obtain a feature map of 2 * 2 * 4. The PassThrough layer itself does not have the ability to learn parameters. As Fig. 2b illustrates, the feature map of the front layer after the feature rearrangement is passed to the next layer, thereby reducing the receptive field of the network and achieving fine-grained features detection.

3 Anchor Boxes for Detection in Specific Areas

For improve the efficiency of the network, we use anchor box to separate the picture into different part and propose the blankets which may possibly contain obstacles before inputting the picture into the network. The anchor box strategy was first proposed in the RPN network of Faster R-CNN and achieved good results [5]. Its essence is the reverse derivation of the idea of Spatial Pyramid Pooling (SPP) [6]. The function of SPP is to resize the input of different sizes into the same size, which

satisfies the adaptability of the network to different size pictures. Anchor box can be understood as a fixed window sampling on a feature map generated after convolution. In the RPN network, each sampling window predicts nine suggested boxes of different sizes and proportions with the center of the window as the center coordinate [7]. Each suggestion box is bound to its own confidence value containing the size of the target possibility. When the target to be detected is not included, the confidence value is 0, and when the target to be detected is included, the confidence value is 1. A suggestion box with a confidence value of 1 is entered into the classification and identification network for further processing.

At the same time, the size of the anchor box in the priori box strategy is different, which helps to detect and identify targets with different scales in the image. At the same time, too many recommendation boxes will increase the running load of the final detection classification network and put forward certain requirements for the operation efficiency of the classification network.

The size and number of the anchor boxes in the RPN network need to be manually set (generally set to 9), which greatly increases the human interference factor and reduces the self-learning effect of the network. Therefore, in order to reduce the interference of human factors in the operation of the network, the prior frame is generated by K-means clustering. The clustering on specific database shows that the number of center points of each object in the image is 5, which can be used to achieve the same effect as the 9 a priori frames of the RPN network. The priori box detects and identifies the object whose center point falls on the mesh.

Based on the anchor box theory, we used the database that collected from the tram's operating to calculate the most 5 possible positions of the obstacles in every frame with the help of K-means clustering (as Fig. 3 illustrates). Then set the anchor boxes at the calculated sites to make sure the network cost more detection in pointed area, which can obviously help to improve efficiency.

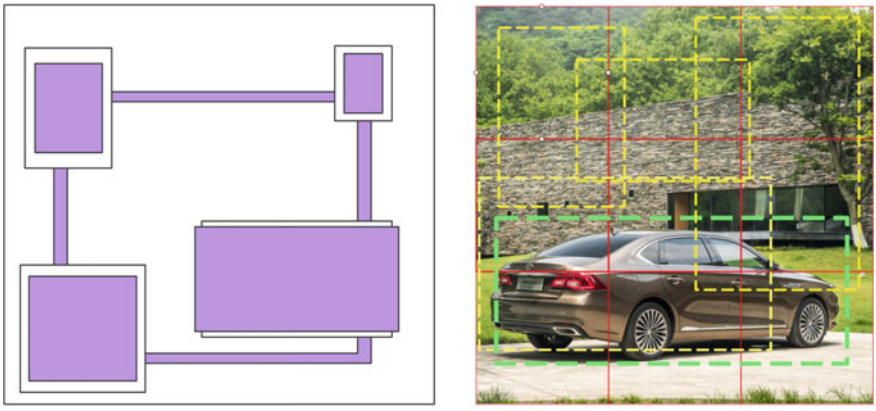


Fig. 3 The illustration about the principle of anchor boxes

4 Batch Normalization for Flexibility in Different Distributions

For training network with great efficiency, we add Batch Normalization (BN) into the network. BN plays an important role to eliminate the disadvantages of low efficiency and poor robustness of network learning due to the different distribution between training data and test data. Meanwhile, for the deep hidden layer, the BN can help to make the input distribution, which is gradually mapped to the nonlinear function and closing to the limit saturation region of the value interval, forcibly pull back to the standard normal distribution with a mean of 0 variance of 1. Then to make the input value of the nonlinear transformation function fall into the area sensitive to the input in order to avoid the gradient disappearing problem (Fig. 4).

Usually, among the trams' operating database, there are some different batch training data features with different distributions (batch gradient decline), the network needs to adapt to different distributions of data in each iteration process, which leads to the inability to use higher learning rate in the training network process and then result in low efficiency of network training [8]. For instance, the network, which is normally trained among the database with the characters that the most obstacles are in the right side, has less accuracy in database with the obstacles in left side. At the same time, the network only learns the feature distribution of the training data, and it is difficult to adapt to the test data with different distribution of the training data features, in which case will result in insufficient robustness of the network. Based on this condition, BN is proposed to solve the situation that the distribution of data in the middle layer changes during the training process. In essence, BN is a

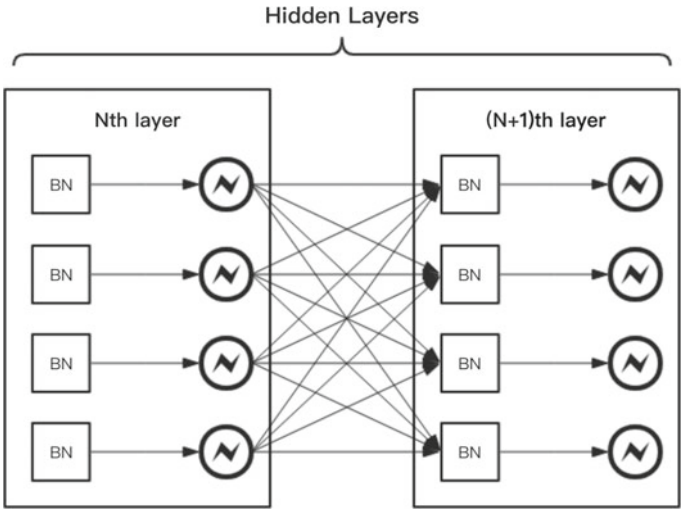


Fig. 4 The structure of batch normalization

learning, parameterized and normalized network layer, which normalizes the output characteristics of the convolutional layer.

However, if only the convolved feature map is normalized, it will affect the characteristics learned by the layer network. Therefore, Batch Normalization contains two learnable parameters γ and β to make the function flexible to different database.

$$y^{(k)} = \gamma^{(k)}\hat{x}^{(k)} + \beta^{(k)} \tag{1}$$

where $\hat{x}^{(k)}$ is the output of the Kth neuron of the layer; $\gamma^{(k)}$ is parameter γ that can be learned in this layer; $\beta^{(k)}$ is parameter β that can be learned in this layer; k is the number of layers. Each neuron in the formula has a pair of learnable parameters γ, β. This way when:

$$\gamma^{(k)} = \sqrt{Var\left(x^k\right)} \tag{2}$$

$$\beta^{(k)} = E\left(x^k\right) \tag{3}$$

The learned feature distribution can be restored. The forward conduction formula for Batch Normalization is:

Mini-batch mean:

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{4}$$

Mini-batch variance:

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_B)^2 \tag{5}$$

Normalize:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \tag{6}$$

Scale and shift:

$$y_i = \gamma\hat{x}_i + \beta \equiv BN_{\gamma,\beta}(x_i) \tag{7}$$

where m is the size of the Mini-batch; x_i is the output of the ith neuron of the layer; ϵ is an error adjustment parameter added to avoid the instability of the numerical calculation to avoid the variance of 0, generally 1e−6.

According to Fig. 5, it is obvious that the network with BN are more likely to extract the characters of pictures in different distribution into a same distribution

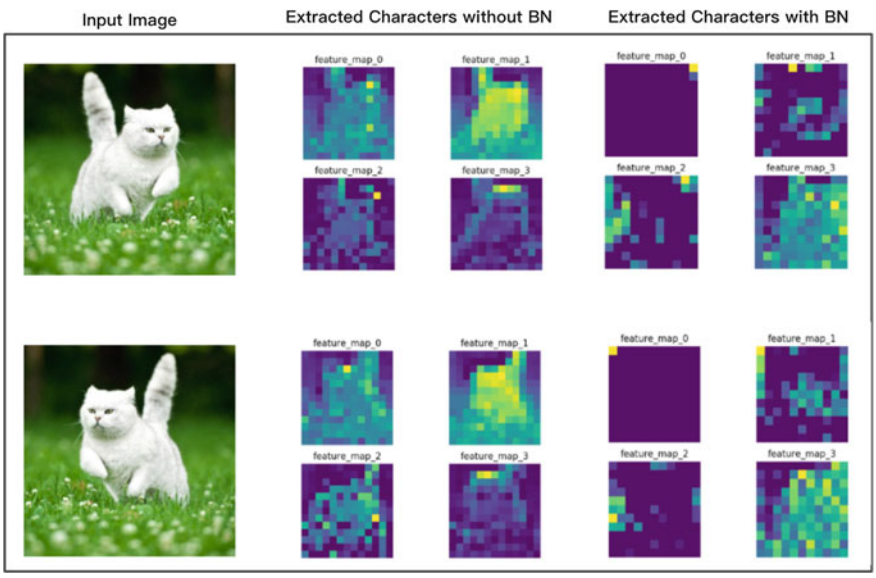


Fig. 5 The extracted characters comparison between network with and without batch normalization

range, which will greatly help to decrease the pressure of the extraction in deep layers and proceed the efficiency to train the network. Additionally, BN can sometimes replace the dropout and make the network more flexible to pictures in different distributions.

5 Network Architectures

The structure of the network is based on the Inception structure which was firstly proposed in the GoogLeNet network [9]. The structure of Inception can be obviously illustrated in Fig. 6. The Inception structure mainly consists of 1 * 1, 3 * 3 and 5 * 5 convolution layers, but the 5 * 5 convolution layer in the Inception structure contains a very large amounts of parameters, that will extremely increase computation [10]. Therefore, the SSD network's idea is referenced in the aspect of saving the optimization network parameters, mainly using 1 * 1 and the 3 * 3 convolutional layer as the image feature extraction layer of the network, in which case can effectively reduce the number of network parameters and save the computational resources occupied by the network. At the same time, the network uses the 2 * 2 Max Pooling Layer in the feature extraction part to reduce the amount of data. In the target classification and recognition part, the global average pooling layer is used to classify the classification results in a one-dimensional array, and to make the final results as formal outputs.

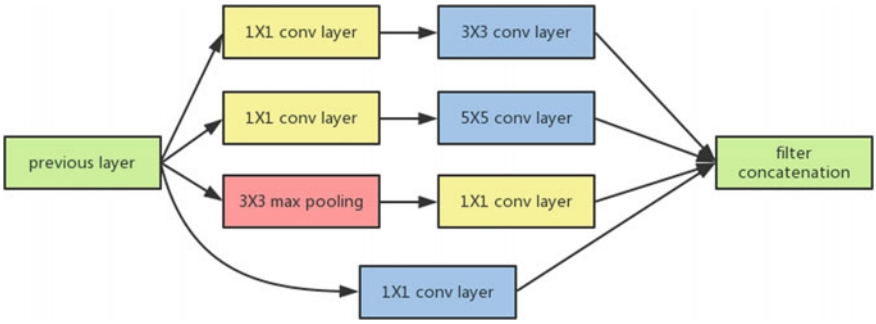


Fig. 6 The original structure of Inception

The concept of a global average pooling layer was first proposed in the Networks in Networks [11]. The working principle refers to the pooling mode of the average pooling layer, but instead of performing the pooling operation based on the size of the core, the entire feature map is averaged. If you input the feature map of S * S * N, the output is a 1 * N one-dimensional vector.

Based on the structure of Inception, the overall network is designed to use 19 convolutional layers, 5 maximum pooling layers, and a global average pooling layer. The feature extraction layer consists of 18 convolution layers and 5 maximum pooling layers. The classification recognition consists of a convolution layer and a global average pooling layer, and the final output is normalized by normalization layer.

The overall network uses the Leaky ReLU function as the activation function, which can effectively alleviate the gradient disappearance problem of the depth network while realizing the sparse activation of the network [12]. At the same time, all normal convolutional layers of the network are added to Batch Normalization, which normalizes the feature distribution acquired by the convolutional layer and improves the robustness of the network network. The Leaky ReLU function is shown as below, where the parameter a is more than 0 less than 1.

$$y(x) = \begin{cases} x, & x \geq 0 \\ ax, & x < 0 \end{cases} \tag{8}$$

In the input stage of the network, the image is resized by Gaussian image pyramid [13]. Referring to the sliding window principle, the K-means clustering is used to generate 5 different sizes and positions with different K-means clusters. An anchor box is designed to detect targets of different sizes at different locations. Each priori box is responsible for predicting the probability of inclusion of the target and the type of the target, and binding the probability value to the a priori box, and outputting is regarded as a candidate suggestion box at the output layer.

The picture that generates the a priori box is input into the network, and after convolution pooling and classification prediction, the prediction result of each candidate suggestion box is output. However, at this time, there are multiple candidate

suggestion boxes to detect the same target, so we use non-maximum suppression [14] to filter out the redundant candidate suggestion boxes.

In the training process, the loss function uses the least square error loss function, and the error back propagation uses the stochastic gradient descent. The loss function refers to the YOLO network on the basis of the minimum mean square error [15]:

$$\begin{aligned} loss_t = & \sum_{i=0}^{W} \sum_{j=0}^{H} \sum_{k=0}^{A} (1_{MaxIoU<Thresh} \lambda_{noobj} * \left(-b_{ijk}^{o}\right)^2 \\ & + 1_{t<12,800} \lambda_{prior} * \sum_{r \in (x,y,w,h)} \left(prior_k^r - b_{ijk}^r\right)^2 \\ & + 1_k^{truth} (\lambda_{coord} * \sum_{r \in (x,y,w,h)} \left(truth^r - b_{ijk}^r\right)^2 \\ & + \lambda_{obj} * \left(IoU_{truth}^k - b_{ijk}^o\right)^2 \\ & + \lambda_{class} * \sum_{c=1}^{C} \left(truth^c - b_{ijk}^c\right)^2) \end{aligned} \tag{9}$$

where w is the width of the feature map; H is the height of the feature map; A is the number of a priori boxes; λ is the weighting factor of each loss part.

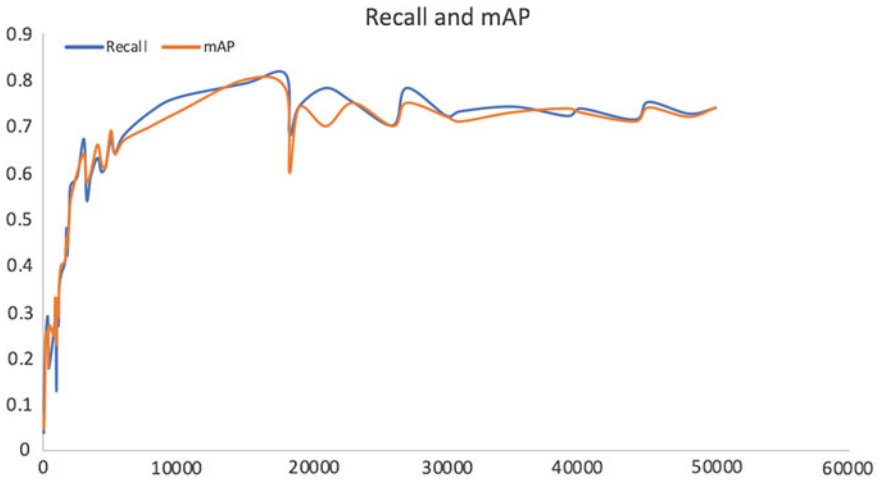
The first loss is used to calculate the confidence error of the background; the second loss is used to calculate the coordinate error of the priori box and the prediction frame; the third loss is calculated by the loss of the prediction frame with the ground truth, including the coordinate error and the classification error and confidence error.

6 Discussion

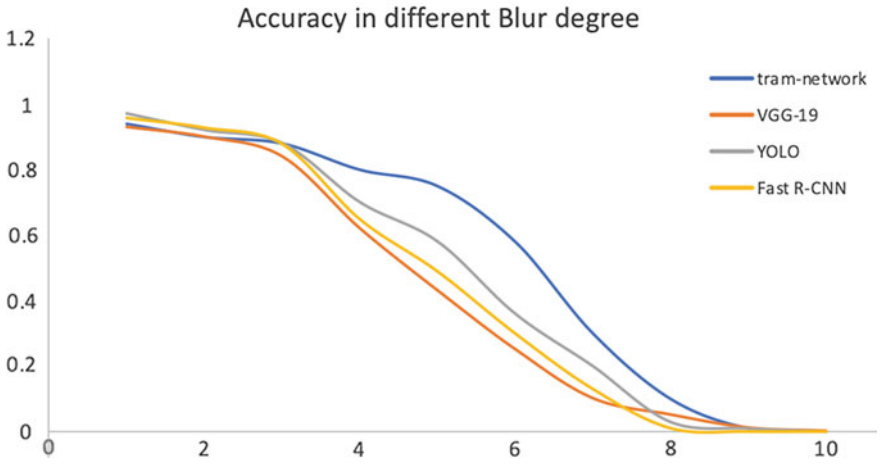
The network's test selects the database containing the 20 types of targets attached to the VOC2007 database, compares the target category position label in the test set with the actual detection target category label, and calculates the recall and mAP indexes of the network while training.

The network has a good detection effect on static images in different sizes and densities. It is also robust to the detection of overlapping targets. At the same time, the network has been tested on different tones and blurred pictures. As can be seen from Fig. 7b, the network has high robustness to target recognition under different definitions and different tonal conditions. We tested the network in different conditions and collected the accuracy rates, which are shown in Table 1. Obviously, the network performs great in normal daytime, while some objections may needed be detected are not that easily to be recognized in some weather conditions.

The network real-time in situ detection is based on the on-board camera to detect and identify the roads in Beijing Datun Road Tunnel and the road near the tunnel. It



(a).The target of Recall & mAP



(b). The accuracy of the network in different blur degree

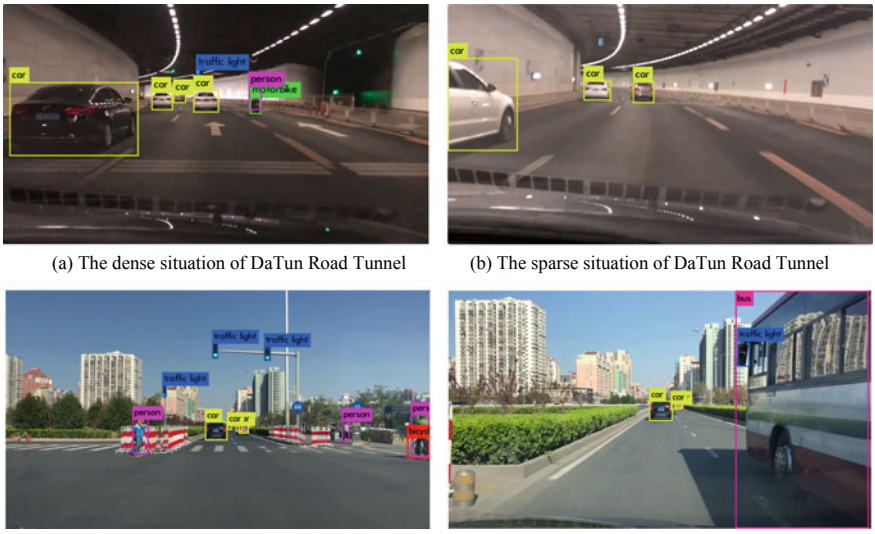
Fig. 7 The evaluation of the network

Table 1 The detection accuracy of different obstacles in different conditions

Objections	Snow (%)	Rain (%)	Night (%)	Sun (%)
Pedestrian	85.2	89.4	81.8	94.7
Car	89.6	84.1	89.3	95.1
Truck	83	80.5	75.4	90.3
Traffic light	67.5	59.4	88.6	85.8
Traffic sign	51.6	55.1	42.8	80.3

is used to test the effect of the network on the in-transit target under different traffic conditions and different lighting conditions. The video partial frame is detected as shown in Fig. 8.

At the same time, after comparison test, the network can realize the real-time detection result of 5 FPS under the GPU with 2 GB GTX750M graphics card. The real-time detection of 3.5 FPS can be maintained under the overheat condition of the graphics card; 25–30 FPS can be realized under the GTX1060 graphics card driven by the GPU. Real-time detection of 30 FPS can maintain real-time performance above 25 FPS under high load and overheat condition of the graphics card. It is worth mentioning that the network also has a certain computing power on the CPU, but because the CPU is not as distributed as the GPU in the calculation unit, the computational efficiency is also reduced. After testing, the network has only FPS of 0.4 on the i7-4850 HQ processor. The FPS is 3.0 on the i7-8700 K processor (Table 2).



(a) The dense situation of DaTun Road Tunnel

(b) The sparse situation of DaTun Road Tunnel

(c) The dense situation of DaTun Road

(d) The sparse situation of DaTun Road

Fig. 8 Real-time detection in route

Table 2 Comparison of network performance index

Network name	Recall	mAP	FPS (2GBGTX750M)	FPS (6GBGTX1060)
Tram-network	73.76	74.3	3.5–5	25–30
YOLO	92.52	78.6	6	45
SSD	87.69	73.8	20	80
Faster R-CNN	95.36	79.4	5	25
ResNet	92.44	79.2	2.5	20

7 Conclusion

The network has a certain good effect in both statistic pics and videos. According to the test in different weather and light conditions, it is obvious that the network is qualified to detect obstacles in complex environments with great efficiency and accuracy. The PassThrough layer helps to extract the characters of tiny objections and make the detection of obstacles in a far distance timely possible. The anchor box provides the possible area of the picture (frame, image) that may contain the obstacles and significantly increase the accuracy and condense the detecting time. The Batch Normalization makes the network more flexible to pictures in different distributions and help to promote the efficiency when training the network.

But there are still some shortcomings in the detection of small objects and the detection accuracy of different distributed objects. It is necessary to continue to optimize the network: optimize the target feature distribution of the training set; increase the training set pair and the number of different distribution samples for a single target; the type of target for streamlining the training set; optimizing the network structure and adding more Passthrough layers for better fine-grained feature detection. The future development of the network can be applied not only to the detection of obstacles in the virtual track, but also to the target detection and identification items in different fields, such as facial expression detection, target segmentation, and species identification.

Acknowledgements This work was supported by the National Key Research and Development Plan (No. 2018YFB1201601-07).

References

1. Currie G, Reynolds J (2011) Managing trams and traffic at intersections with hook turns. Transp Res Record: J Transp Res Board 2219:10–19
2. Samala RK, Chan HP, Hadjiiski LM et al (2017) Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms. Phys Med Biol 62(23)
3. Hertel L, Barth E, Käster T et al (2017) Deep convolutional neural networks as generic feature extractors
4. Shelhamer E, Long J, Darrell T (2014) Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal Mach Intell 39(4):640–651
5. Ren S, He K, Girshick R et al (2015) Faster R-CNN: towards real-time object detection with region proposal networks. In: International conference on neural information processing systems. MIT Press, Cambridge, pp 91–99
6. He K, Zhang X, Ren S et al (2014) Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell 37(9):1904–1916
7. Girshick R (2015) Fast R-CNN. Computer Science
8. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International conference on international conference on machine learning

9. Szegedy C, Liu W, Jia Y et al (2015) Going deeper with convolutions. In: Computer vision and pattern recognition. IEEE, New York, pp 1–9
10. Szegedy C, Vanhoucke V, Ioffe S et al (2015) Rethinking the inception architecture for computer vision. Comput Sci 2818–2826
11. Giusti A, Dan C C, Masci J et al (2013) Fast image scanning with deep max-pooling convolutional neural networks
12. Schmidt-Hieber J (2018) Nonparametric regression using deep neural networks with ReLU activation function
13. Olkkonen H, Pesola P (1996) Gaussian pyramid wavelet transform for multiresolution analysis of images. Graphical Models Image Process 58(4):394–398
14. Neubeck A, Gool LV (2006) Efficient non-maximum suppression. In: International conference on pattern recognition. IEEE, New York, pp 850–855
15. Redmon J, Farhadi A (2016) YOLO9000: Better, Faster, Stronger, 6517–6525

Prediction of Failure Rate of Metro Vehicle Bogie Based on Neural Network

Xiuqi Wang, Yong Qin, Yong Fu and Meng Ye

Abstract Metro train bogie system is located between the car body and the track, which is one of the key subsystems to ensure the safety of train operation. As a complex system, bogie system is composed of many components, once the failure happened, it would impact the normal operation of the whole train. In order to better predict the failure rate of bogie system, radial basis function (RBF) neural network is introduced to predict the failure rate of the whole system through the fault data of bogie components, and genetic algorithm is used to optimize the model. Experimental results showed that the proposed method can accurately predict the bogie failure rate, and can be used as a system-level reliability prediction method, providing a data basis for later system improvement and optimization.

Keywords Bogie · Genetic algorithm · RBF neural network · Prediction of failure rate

1 Introduction

As the key system of the train, the bogie is responsible for the functions of traction, steering and braking. Whether it regularly works directly affects the safe operation of the train. Therefore, the reliability analysis and failure rate prediction of bogie system have very important practical significance for the safety and stability of train operation. According to the structure and function, it can be divided into the following eight kinds of equipment: frame, wheelset, axle box, primary spring suspension device, secondary spring suspension device, driving device, basic brake device and central traction connection device. The operation safety and stability of the bogie

X. Wang · Y. Qin (✉) · Y. Fu · M. Ye
National Key Laboratory of Rail Control and Safety of Beijing Jiaotong University, Beijing 100044, China
e-mail: yqin@bjtu.edu.cn

Y. Qin
Beijing Engineering Research Center of Urban Traffic Information Intelligent Sensing and Service Technologies, Beijing 100044, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_83

are determined by the interrelation of various kinds of equipment. Therefore, the failure rate of a single equipment can not represent the whole system. In order to achieve a comprehensive failure rate prediction, it is necessary to introduce a method to connect each equipment with the whole bogie and build a "equipment-system correlation model" to realize the failure rate prediction of the whole system.

The commonly used methods of train system reliability prediction include FMECA, fault tree analysis, Markov model, complex network analysis and so on. Xia Jun found out the key components and weaknesses of the system through FMECA, and predicted the reliability of the train door system by combining the interface development tools and the programming language [1], but the fault data of the equipment do not keep the linear distribution in accordance with the time sequence, which makes the accuracy of this method low; Hu-chuan et al. started with the causal relationship of the fault, refined the whole to the part step by step, and established the fault tree, so as to get the causes of the system faults, and provided data support for the formulation of maintenance schemes [2]. In order to overcome the polymorphism of events and the logical uncertainty between events, Wang Heng-liang et al. introduced Bayesian network, which improved the accuracy of reliability analysis of running-gear [3], but the task of constructing fault tree was heavy and difficult. Meng Ling-hui et al. established Markov model according to the fault mechanism and train system state, analyzed key components and established fault model to characterize their failure rate, so as to obtain the changing trend of train reliability and realize vehicle life prediction [4]. However, Markov model requires plenty of data, which are both fault data and maintenance data. Qin Yong et al. considered the complexity of the structure and the diversity of fault modes of the train system. Combining the cause and effect mechanism of fault with the structure of train, they established a multi-network to analyze the reliability of train and came true more complete and accurate prediction [5]. However, to build a complex network, it is necessary to know the fault mechanism of each fault mode and the logical functional relationship between the components of train.

Metro bogie system is a complex mechatronic system, which is mainly manifested in: (1) there are many components with strong related affection, and the structure between each other is complex; (2) the impact of each component on the whole bogie is different; (3) the failure modes are complex and diverse, and the failure data of each component have highly nonlinear characteristics. According to the data acquisition of bogie system at present, there are only fault data but no maintenance data, and the internal structure of the system and the correlation between components cannot be accurately described, so the application of the above method is difficult.

In this paper, the reliability prediction is realized by the neural network model. On the one hand, the prediction error of the neural network is smaller than that of the traditional mathematical model; on the other hand, the neural network can be modified by learning when the data is incomplete; besides, the neural network can realize a highly nonlinear mapping from input to output, which is suitable for the data with nonlinear changes [6].

At present, many studies have used neural networks: Zhou et al. [7] established RBF neural networks based on the operating state data of metro trains to predict

the failure location; Song et al. [8] presented an on-line fault detection diagnosis method of two channel speed sensor based on RBF neural network to improve the installation space of onboard equipment and ensure the safe operation of urban rail train; Li et al. [6] established BP neural networks to predict metro train reliability; Yin et al. [9] proposed a model for axle box failure rate prediction based on PSO-BP neural network. The simulation results showed that the effect of the model is pretty well.

Considering that BP neural network is global optimum approach and rates of convergence is slow, this paper used RBF neural network. RBF network has the advantages of fast learning convergence rate, strong classification ability, the beat uniform approximation, not easily falling into local minimum, and strong mapping function between input and output compared with other networks [10].

In this paper, the RBF neural network is used as a bridge, and the failure rates of eight types of equipment and the whole system are taken as input and output respectively. The "equipment-system" correlation prediction model is established to predict the failure rates of bogie system and then the model is optimized by using genetic algorithm.

2 RBF Neural Network

2.1 Definition

A RBF neural network can be divided into three layers: an input layer which can accept data from user, a nonlinear hidden layer bringing nonlinear response to output layer, and an output layer which is used to output data from the hidden layer, these data have be linear weighted combined. The structure of a RBF neural network with n inputs, p output is given in Fig. 1 [11]. The function of hidden layer is radial basis functions which has many forms and Gaussian activation function is one of the most used functions.

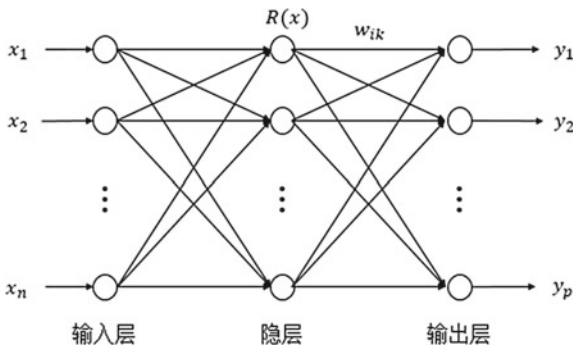


Fig. 1 The structure of RBF neural network

$$R_i(x) = \exp\left(-\frac{||x - c_i||}{2\sigma_i^2}\right) \quad i = 1, 2, \ldots, m$$

where, input $x = [x_1, x_1, \ldots, x_n]$, c_i and σ_i are the center and width of the corresponding membership function, respectively. $||x - c_i||$ is the distance between x and c_i, m is the number of hidden layer nodes.

When Gaussian activation function is chosen, $x \rightarrow R_i(x)$ (from input layer to hidden layer) is linear mapping, $R_i(x) \rightarrow y_k$ (from input layer to hidden layer) is nonlinear mapping. The input vector is denoted as $x = [x_1, x_1, \ldots, x_n]$, the output vector is denoted as $y = \left[y_1, y_1, \ldots, y_p\right]$, therefore the kth neuron in output layer is:

$$\widehat{y_k} = \sum_{i=1}^{m} \omega_{ik} R_i(x) + b_k \quad k = 1, 2, \ldots, p$$

where, n is the number of input layer nodes, m is the number of hidden layer nodes, p is the number of output layer nodes, ω_{ik} is the weight between the ith neuron in hidden layer and the kth neuron in output layer, b_k is threshold.

So, as soon as c_i and ω_{ik} is determined, it can be used to find the output value corresponding to the network when given an input [12].

2.2 Shortcomings of RBF Neural Network

Since its application, RBF neural network has been widely recognized for its nonlinear approximation ability, autonomous training function and learning convergence speed, but there are still some problems. The biggest disadvantage is the parameter selection. The improper selection of center and width of the corresponding membership function and weight will directly affect the output of network and reduce the speed and accuracy of prediction. For example, there is a certain randomness in the selection of expansion constant of neural network, which may affect the approximation ability of function, and then lead to network diffusion [13].

To solve the above problems, this paper introduced genetic algorithm to optimize network parameters. Genetic algorithm is different from the traditional method which uses inherent models or formulas to solve directly, but continuously chooses, crosses and varies among all known feasible solutions, and finally produces the optimal solution.

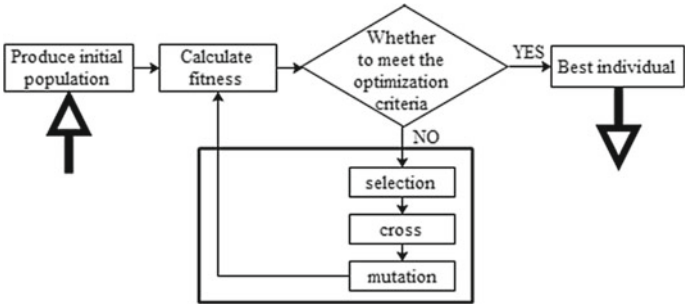


Fig. 2 Optimization flow chart of genetic algorithm

2.3 Genetic Algorithm

Genetic algorithm is one of the common methods to optimize the parameters of neural network. The essence of genetic algorithm is to find the optimal solution by calculating the probability. In the process of searching, it can adjust the direction adaptively and avoid falling into the local minimum. It has good self-organization and randomness. It can carry out parallel operation and search for the optimal solution iteratively. Moreover, genetic algorithm is a combination of coding parameters and has strong generality [14] (Fig. 2).

3 Cases of Bogie System Failure Rate Prediction

3.1 Modeling Approach

Statistical analysis of bogie operation failure data showed that the bogie's failure is determined by eight types of equipment: frame, wheelset, axle box, primary spring suspension device, secondary spring suspension device, driving device, basic brake device and central traction connection device. The purpose of the model is to predict the failure rate of the whole system through the failure rate of the above eight kinds of equipment. Therefore, the RBF neural network model is established with the failure data of the eight kinds of equipment as input value, that is, the number of nodes in the input layer is eight, and the failure data of the system as output value which means the node in the output layer is one.

The fault data of eight kinds of equipment are respectively the number of faults of bogie equipment recorded during train operation. The fault data of the whole bogie system can be obtained by analyzing the severity of the faults of each equipment. The faults of equipment can be divided into two categories according to the severity of the consequences. The first type is the faults that can cause the damage of bogie

working characteristics and incompleteness which will lead to delays in train departure, interruptions in operation, and evacuation of passengers. The second type is all other faults, which will not affect the normal operation of bogies. The sum of the first type failures of all the equipment is the number of failures of the whole bogie system [15].

3.2 Modeling

(1) Data preprocessing

Taking a certain type of metro train bogie system as an example, the fault data recorded during the operation period are counted. According to the fault statistics principle mentioned above, the fault data of eight key equipment are obtained. Taking fifteen days as a cycle, a total of 102 groups are obtained. Then the failure data are transformed into failure rate by formula 3 [16], and the results are shown in Table 1.

$$\lambda_i = \frac{N_i}{T}$$

where, λ_i is the failure rate of the ith cycle, N_i is the total number of failures of the ith cycle, T is the total time of the ith cycle.

(2) Establish predictive models

- Read data and divide into groups: The processed data are divided into two groups: 1–82 groups as training samples, 83–102 groups as test samples;
- Create a $newrbe$ neural network:

$$net = (newrbeP\_train, T\_train, 0.0756)$$

- Simulation and performance evaluation: select Mean Squared Error (MSE) and R-Square (R^2) as the criteria for network performance evaluation which

Table 1 Data grouping of failure

二系悬挂装置	构架	轮对	轴箱装置	中央牵引连接装置	驱动装置	制动装置	一系悬挂装置	转向架
0	0	0	0	0.067	0.067	0.067	0	0.133
0	0	0	0	0	0	0	0	0
0	0.067	0.2	0	0	0.067	0	0.067	0.2
0	0	0.067	0	0	0	0	0	0.067
0	0	0	0	0.067	0	0	0	0.067
0.067	0	0	0.067	0.067	0	0.067	0.067	0.2

means the accuracy of the network is higher when MSE is lower and R^2 is bigger.
- Output results: the comparison chart of prediction results, MSE, R^2 and correlation coefficient.

(3) Genetic algorithm optimization

In order to improve RBF neural network and reduce its prediction error, the extended constant of RBF neural network is optimized by genetic algorithm.

In the use of genetic algorithm, four parameters need to be set in advance: population size, iteration times, crossover probability and mutation probability. These four parameters have a certain impact on the results and efficiency of the algorithm, but at present there is no reasonable theory for setting them. It is often need to determine a reasonable range of parameters after many experiments and comparisons.

- Encoding: real coding of expansion constants;
- Initial population: Experiments show that the optimal initial population is 20–70. In this network, the initial population is 20.
- Fitness function evaluation: The purpose of fitness function is to "preserve the good and remove the bad". The individuals with large fitness are retained for better inheritance. In neural networks, the MSE is usually used as the fitness function.

$$f = sum\big((T\_sim - T\_test).^2\big)/length(T\_sim)$$

- Selection operation: This model chooses roulette selection. According to the ratio of the fitness of each individual to the total fitness of the population, higher ratio has higher possibility to enter the next generation.
- Crossover operation: Use arithmetic cross and the probability of crossover is 0.8.
- Mutation operation: Use uniform mutation and the probability of mutation is 0.4.
- End of Evolution: Set the number of iterations to 200. When the set value is reached, the evolution stops and the optimization results are saved.
- Create a *newrbe* neural network:

$$net = (newrbeP\_train, T\_train, bestr)$$

3.3 Simulation and Analysis

After simulation, the results are shown in the following figures:

(1) Results of RBF neural network
(2) Results of GA-RBF neural network

It can be seen from Figs. 3, 4, 5 and 6 that MSE and R^2 of the RBF neural network test data are 0.0589 and 0.3017. The network output value and the actual

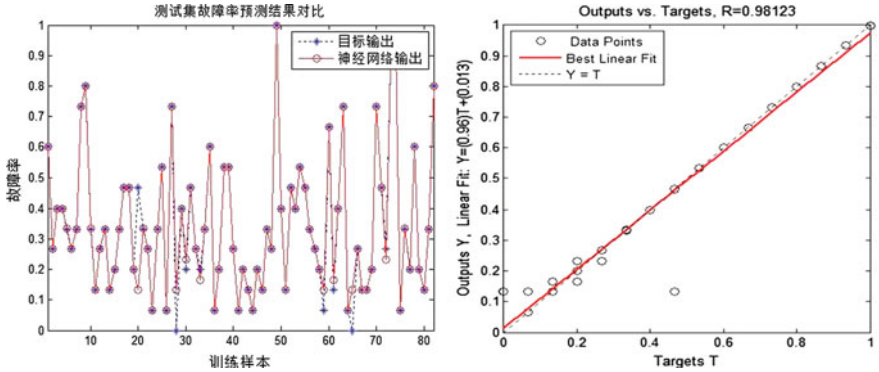


Fig. 3 Contrast between RBF neural network training output and target output

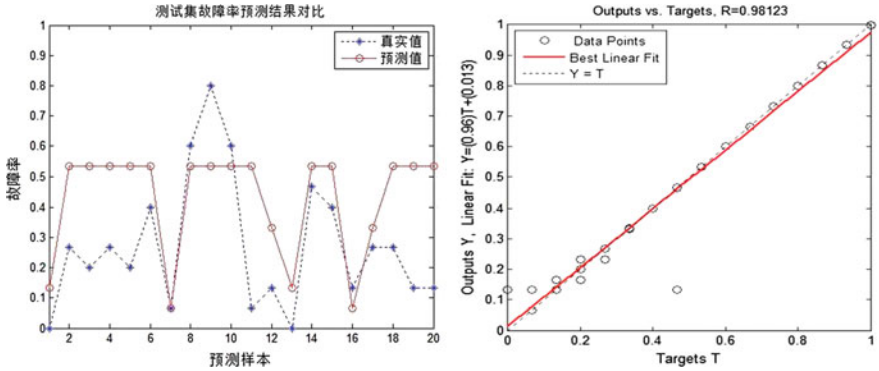


Fig. 4 Contrast between RBF neural network test output and target output

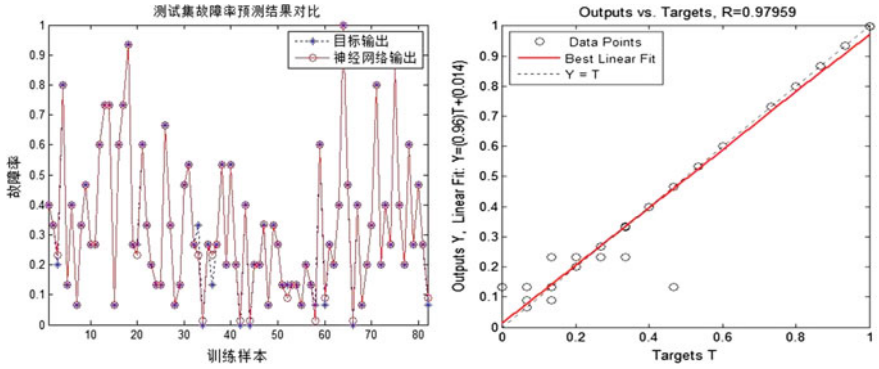


Fig. 5 Contrast between GA-RBF neural network training output and target output

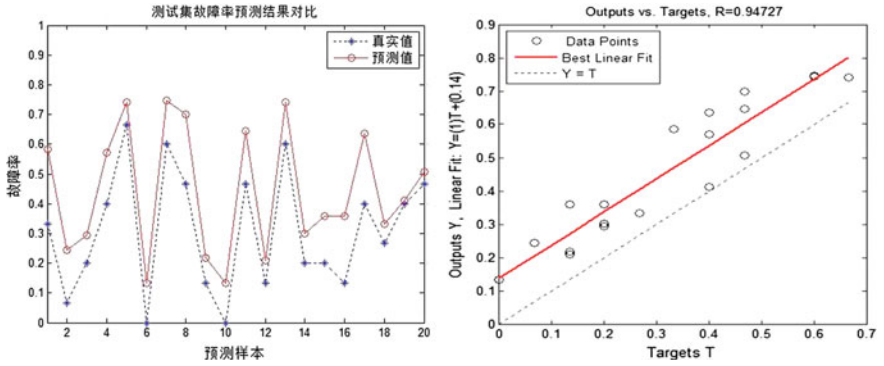


Fig. 6 Contrast between GA-RBF neural network test output and target output

output value of the training data are almost completely coincident. The correlation coefficient is 0.981, which is pretty well. However, there is almost no coincidence between network output value and the actual output value of the test data, and the correlation coefficient is also relatively low, only 0.549, which means the correlation is poor and the prediction accuracy need be improved. It indicates that the RBF neural network is not suitable as a "device-system-associated failure rate model" to achieve the reliability prediction of the bogie system level. On the contrary, the's MSE and R^2 of test data are 0.0231 and 0.8973, the correlation coefficient between the network output value and the actual output value of the training data is 0.980, which is relatively good, and the correlation coefficient between the network output value and the actual output value of the test data is 0.947, which is pretty nice too, indicating that the GA-RBF neural network can be used as a "equipment-system" correlation prediction model to achieve system-level reliability prediction.

4 Conclusion

Firstly, we have introduced the RBF neural network and analyzed its shortcomings. On this basis, we have proposed using genetic algorithm to optimize parameters in order to obtain the GA-RBF neural network model. Through MATLAB, we have realized the related failure rate prediction model of bogie from equipment to system. The simulation results have shown that GA-RBF neural network model has high accuracy and can better realize the failure rate prediction of bogie system.

However, due to time and capacity constraints, this paper only considered the impact of internal equipment on the system failure rate, and only used genetic algorithm to optimize. In the follow-up study, other external factors, such as operating environment, can be added, as well as more algorithms can be use to improve the network structure so that we could get the best prediction model.

Acknowledgements The authors gratefully acknowledge the financial supports for this research from National Natural Science Foundation of China (61833002).

References

1. Xia J (2014) Research on reliability analysis and application of metro door system. Nanjing University of Science and Technology, Nanjing
2. Hu C, Yao J (2012) Reliability analysis for electric multiple units based on fault tree Monte Carlo method. China Railway Sci z1:52–59
3. Wang H, Lu Z, Zhang B (2012) Analysis method for the operational reliability of EMU running gear based on fault tree and Bayesian network. China Railway Sci z1:60–64
4. Meng L, Liu Z, Diao L et al (2016) Reliability evaluation of high-speed train traction transmission system based on Markov model. J China Railway Soc 8:23–27
5. Qin Y, Fu Y, Li W et al (2018) Operational safety and reliability assessment of high speed train with intuitionistic fuzzy set and VIKOR method. J Beijing Univ Technol 44(1):112–119
6. Li L, Cheng X, Qin Y et al (2013) Reliability prediction of urban rail transit vehicle based on BP neural network. J Central South Univ (Science and Technology) 1:42–46
7. Zhou Q, Deng Y, Chen J (2010) The RBFNN application in fault diagnosis for the subway train. Comput CD Softw Appl 11:71–72
8. Song Y, Zhu M (2014) Subway sensor fault diagnosis based on radial neural network. Urban Mass Transit 5:94–97, 101
9. Yin H, Wang K, Zhang T (2015) Fault prediction based on PSO-BP neural network about wheel and Axle Box of Bogie in urban rail train. Complex Syst Complexity Sci 4:97–103
10. Hartman E, Keeler JD, Kowalski JM (2008) Layered neural networks with Gaussian hidden units as universal approximations. Neural Comput 2(2):210–215
11. Zhou P (2013) Design and application of neural network based on Matlab. Tsinghua University Press, Beijing
12. Yu J (2012) Reliability analysis and application research of key system of metro vehicles. Beijing Jiaotong University, Beijing
13. Zeng Y (2015) RBF natural network based on genetic algorithm used in maximum power point tracking of photovoltaic system. Hunan University of Technology, Zhuzhou
14. Chen A (2007) Research on data prediction method based on BP and RBF neural networks. Central South University, Nanjing
15. Chen M (2006) Study of submarine's displacement and principal dimensions by using GA based optimum RBF neural network. Huazhong University of Science and Technology, Wuhan
16. Shi J (2014) Study of Bogie failure prediction and maintenance of urban rail train based on reliability analysis. Beijing Jiaotong University, Beijing

Vehicle Risk Analysis and En-route Speed Warning Research Based on Traffic Environment

Jian Xiong, Yan-li Bao, Zhou-jin Pan and Yi-fan Dai

Abstract In this paper, we established a method to assess the risk of traffic environment and proposed a vehicle speed early warning model that is related to the traffic environmental risk index. This risk assessment method of traffic environment takes into account three risk factors: the risk exposure, the probability of accident occurrence and the severity of accident. Simultaneously, two dynamic risk factors of operating speed and real-time speed are also introduced. The vehicle speed warning model is correlated with the Traffic Environment Risk Index. The establishment of Early Warning Vehicle Speed and Vehicle Speed Early Warning System are based on the Evaluation Index of Operating Vehicle Speed and Velocity Gradient. The traffic environment risk assessment method and speed early warning model are validated by using the accident data and the section speed data of Kunshi Expressway, respectively. The Spielman correlation coefficient between the risk grade and the number of road accidents is determined from our experiments to be 0.7109. The average value of speed gradient of early warning vehicle speed is less than that of running vehicle speed gradient. The accident rate of early warning speed is lower than that of running speed in the same section.

Keywords Traffic engineering · In transit vehicle speed warning · Road latency risk · Risk assessment model · Velocity gradient

1 Introduction

In recent years, investment and construction of expressways and urban roads have been intensified in China. The mileage of expressway construction is increasing. According to data released by the National Bureau of Statistics, by the end of 2017, the expressway network had reached 477.36 million kilometers, of which 136.5

J. Xiong (✉) · Y. Bao · Z. Pan · Y. Dai
School of Transportation Engineering, Kunming University of Science and Technology, Kunming 650000, China
e-mail: xjelox@qq.com

Suzhou Automotive Research Institute, Tsinghua University, Suzhou 215000, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_84

million kilometers were expressways. A large number of studies have shown that among the factors leading to traffic accidents, more than 23% are related to the traffic environment [1]. Road conditions and traffic environment are important factors to induce traffic accidents.

At present, the global research on vehicle early warning mainly focuses on the early warning of vehicle active safety. Cui H. Y., Xin S. X. and Feng X. C. analyzed the impact of car interior environment on driving safety [2]. The theory of vehicle interior environment for comprehensive evaluation of traffic safety is constructed. On this basis, a set of on-board early warning system is designed, which can real-time detect, evaluate, display, locate, warn and transmit the environmental elements of the vehicle. Shen B., Zhang Z. and Liu H., et al. proposed an early warning system for collecting real-time vehicle data through roadside sensors and transmitting information to drivers in time through roadside warning lights [3]. In this paper, the principle of alarm system is studied. The accuracy of speed prediction model of early warning system is verified by driving simulation experiment.

Domestic and foreign research on expressway traffic environmental risk mainly focuses on the mechanism of the impact of expressway traffic environment on traffic safety, etc. [4]. In 2018, Wang J. J., Cao X. D. and Yang Y. F., et al. first used the combination method of CRITIC method and rank-sum ratio method of weight calculation to establish a two-dimensional traffic safety risk index system of mountainous road and risk isolation model of input and output, and to classify and evaluate the road risk degree [5]. In 2016, Yang C. F., Gao H. N., Sun J. S., et al. establishes an evaluation index system of road safety grade considering road route, road condition, safety facilities and traffic environment [6]. The model combines subjective and objective weights and is optimized by consistency checking and improved grey relational combination weighting method. Hao Q. Y., Xiong J., et al. determined the evaluation index and method of latent risk of road traffic environment based on the road traffic environment latent risk theory [7].

The above research results promote the development of traffic safety and vehicle early warning. However, in the process of evaluation and vehicle early warning, the qualitative research on traffic safety is limited, and the quantitative research on traffic safety or risk is insufficient. In this paper, an on-road vehicle speed early warning method based on traffic environmental risk index is proposed. It makes a theoretical and quantitative analysis of road traffic environment risk and running vehicle. According to the evaluation theory, real-time vehicle speed warning is carried out for on-the-road vehicles. It proves that the early warning speed is more stable, and it can reduce the accident rate and improve the traffic safety.

2 Modeling of Traffic Environmental Risk Index

2.1 Traffic Environmental Risk Model

Traffic environmental risk refers to the risk caused by potential unsafe risk elements existing in road conditions or traffic environment in the traffic system [8, 9]. This paper focuses on three basic factors that describe expressway safety: the risk exposure, the probability of accident occurrence and the severity of accident [10]. The definition is as follows:

Traffic Environmental Risk = Risk Exposure * Probability of Accident Occurrence * Severity of Accident.

The Risk Exposure is related to traffic flow. The degree of Risk Exposure increases with the increase of traffic flow. Risk Exposure of a expressway can be expressed by Formula (1). The risk exposure level in the formula is divided into 0–5 grades. The higher the score, the higher the risk exposure level.

$$E_i = \left(\frac{V_i}{V_{\max}} \right) \times 5.0 \tag{1}$$

where

E_i Risk Exposure Degree of a expressway section unit,
$V_{\max}$ maximum volume on a expressway,
V_i the volume on a expressway section unit.

As can be seen from Table 1, The target of the probability of accident occurrence grade is P_i; the standard of accident occurrence grade are Road Condition (R), Roadside Environment (E), Transport Facilities (F), Climatic Conditions (C); the index of accident occurrence grade is the information contained in risks of various macro factors (r_i, e_i, f_i, c_i). Constructing judgment matrix by expert consultation and calculating weight coefficient. Then the probability of accident occurrence is established as shown in formula (2):

$$P_i = W_1 \times R + W_2 \times E + W_3 \times F + W_4 \times C \tag{2}$$

where

P_i the probability of accident occurrence of a expressway section unit,
R road condition risk, $R \subset (r_1, r_2, r_3, r_4, r_5, r_6, r_7)$,
E roadside environment risk, $E \subset (e_1, e_2, e_3, e_4, e_5)$,
F transport facilities risk, $F \subset (f_1, f_2, f_3, f_4)$,
C climatic conditions risk, $C \subset (c_1, c_2, c_3)$,
$W_{j(j=1,2,3,4)}$ weights for various types of risks.

Table 1 Factors of accident possible

Target layer	Standard layer	Index layer	Partition criterion
Probability of accident occurrence	Road condition (R)	Curve (r_1)	$R \leq 1000$ m
		Hill (r_2)	$Slope \geq 3\%$
		Lane change (r_3)	Increase/decrease in number of lanes
		Insufficient sight distance (r_4)	Road space sight distance does not meet the requirements of operating speed
		Bridge/tunnel/culvert (r_5)	100–300 m before entering tunnel entrance/bridge starting point; Location to 100–300 m after exit of tunnel entrance/bridge terminal
		Long downhill without truck escape ramp (r_6)	Long downhill without truck escape ramp
		No signal control intersection (r_7)	No signal control at intersection
	Roadside environment (E)	Toll station (e_1)	The road through the toll station
		Service areas (e_2)	The road through the service area
		Parking area (e_3)	The road through the parking area
		Sightseeing stand (e_4)	The road through the sightseeing stand
		Expressway ramp (e_5)	The road through the expressway ramp
	Transport facilities (F)	No barrier (f_1)	No barrier on the roadside
		No sight guidance (f_2)	No sight guidance on the road
		No isolation facilities (f_3)	No isolation facilities on the road
		No anti glare (f_4)	No anti glare on the road
	Climatic conditions (C)	Rain (c_1)	$Precipitation \geq 25$ mm/d
		Snow (c_2)	$Snowfall \geq 0.1$ mm/d
		Fog (c_3)	$Visibility \leq 500$ m

A questionnaire survey was conducted among 30 experts in the transport and intelligent network automotive industry. The weights of the factors are obtained as shown in Table 2:

R, E, F, C is determined by the weight of various kinds of risk information. Among them, The factors with the greatest weight are level 5 risk; the factors with the least weight are level 1 risk; and the weights of intermediate factors are determined by interpolation. If the above information does not exist, the risk value is 0. The Meaning of Risk Level: Level 0 is risk-free; Level 1 is low risk, $S = 1$; Level 5 is high risk, $S = 5$; The median value is between the two. According to the above

Table 2 Weight of layer factors

Target layer	Standard layer (Macro factors)		Index layer (Microcosmic factors)		Comprehensive weight coefficient
	Macro factors	Weights	Microcosmic factors	Weights	
Probability of accident occurrence (P_i)	Road condition (R)	0.307	Hill (r_2)	0.154	0.0473
			Insufficient sight distance (r_4)	0.150	0.0461
			Curve (r_1)	0.149	0.0457
			Lane change (r_3)	0.148	0.0454
			Long downhill without Truck escape ramp (r_6)	0.139	0.0427
			No signal control intersection (r_7)	0.132	0.0405
			Bridge/tunnel/culvert (r_5)	0.128	0.0393
	Roadside environment (E)	0.273	Expressway ramp (e_5)	0.230	0.0628
			Toll station (e_1)	0.221	0.0603
			Service areas (e_2)	0.193	0.0527
			Parking area (e_3)	0.183	0.0500
			Sightseeing stand (e_4)	0.173	0.0472
	Transport facilities (F)	0.210	No barrier (f_1)	0.294	0.0617
			No isolation facilities (f_3)	0.262	0.0550
			No sight guidance (f_2)	0.243	0.0510
			No anti glare (f_4)	0.201	0.0441
	Climatic conditions (C)	0.210	Snow (c_2)	0.428	0.0899
			Rain (c_1)	0.293	0.0615
			Fog (c_3)	0.279	0.0586

definition, let S be the risk level of a certain kind of factor, P be the weight of the risk factor, $P_{\max}$ be the max weight of the risk factor, $P_{\min}$ be the min weight of the risk factor, The Formulas (3) and (4) are as follows:

$$\frac{S-1}{P-P_{\min}} = \frac{1-5}{P_{\min}-P_{\max}} \tag{3}$$

$$S = 1 + \frac{4}{P_{\max}-P_{\min}}(P-P_{\min}) \tag{4}$$

S is a value between 0 and 5. Let $S_{ij(j=1,2,3,4)}$ are risk factor rank among four risk factors of Road Condition, Roadside Environment, Transport Facilities, Climatic Conditions in section unit i, the Formula (6) of the probability of accident occurrence are as follows:

$$P_i = \sum_{j=1}^{4} W_j S_{ij} \tag{5}$$

where

the value of the probability of accident occurrence (P_i) of section unit i is (0, 5).

Maximum restricted speed of a expressway is used to quantitatively calculate the Severity of Accident [10]. The severity of the accident level is divided into 0 to 5. Define the severity of accident as C_i:

$$C_i = \left(\frac{PS_i}{PS_{\max}}\right) \times 5.0 \tag{6}$$

where

PS_i is the restricted speed of a section unit in a expressway,
$PS_{\max}$ is the max restricted speed of a expressway.

For a expressway, the formula for calculating the grade of Traffic Environmental Risk Model is as follows:

$$TERM = E_i \times P_i \times C_i \tag{7}$$

where

E_i is the grade of the Risk Exposure,
P_i is the grade of the Probability of Accident Occurrence,
C_i is the grade of the Severity of Accident.

2.2 Dynamic Risk Factors

Traffic Environmental Risk Model is the Prerequisite for Accidents. Only when the speed of the vehicle reaches a certain level can it bring risks to driving. Operating speed refers to the 85th percentile of the vehicle speed measured at the characteristic points of the road section when the traffic is in a free flow state and the weather conditions are good. Operating speed takes into account the traffic psychological requirements of most drivers. Taking the actual running speed of the vehicle as the linear design speed can effectively ensure the coordination between the route design elements and driving behavior.

Therefore, the dynamic risk model considers the risk effect of operating speed. The consistency and coordination of driving speed can be reflected to a certain extent by using the evaluation method of operating speed coordination. There is a direct relationship between the risk of vehicle driving and braking distance. According to the basic motion model of vehicle, When the vehicle travels at speed v, the reaction time is neglected and the braking distance in emergency braking is $L = \frac{v^2}{2a}$. It can be seen that the braking distance is proportional to the square of the driving speed. Therefore, it can be considered that the driving risk is directly proportional to the square of the vehicle speed. Dynamic risk factors can be expressed as:

$$DRF = \left(\frac{v_i}{v_d}\right)^2 \tag{8}$$

where

v_i Real-time Vehicle Speed,
v_d Operating Speed.

2.3 Traffic Environmental Risk Index

Traffic Environmental Risk Index takes traffic environmental risk model and dynamic risk factors into account. Its formula is as follows:

$$TERI = TERM \times DRF = E_i \times P_{\mathrm{i}} \times C_i \times \left(\frac{v_i}{v_d}\right)^2 \tag{9}$$

where

TERI Traffic Environmental Risk Index,
TERM Traffic Environmental Risk Model,
DRF Dynamic Risk Factors,
E_i Risk Exposure,
P_i Probability of Accident Occurrence,

C_i Severity of Accident,
v_i Real-time Vehicle Speed,
v_d Operating Speed.

3 Verification of Traffic Environmental Risk Index

3.1 Road Selection

Select Kunming-Shilin Expressway (hereinafter referred to as Kunshi Expressway). The expressway is 78.08 km long. The Kunshi Expressway is divided into 16 units every 5 km section. Calculate the risk exposure, the probability of accident occurrence and the severity of accident for each unit and computation of Traffic Environmental Risk Index. At the same time, the number of accidents in each unit is counted, and the relationship between the two groups of data of risk assessment results and the number of accidents is analyzed. Computing Spielman correlation of two sets of data. According to statistics, the flow rate of each unit of Kunshi high-speed is similar, so the definition in this experiment assumes that the risk exposure degree of each unit is the same.

3.2 Model Application

In order to verify the rationality of the Traffic Environmental Risk Index, the risk grade of the section is compared with the 5-year accident number of the section. Considering that the Exposure Risk of this section is the same, neglecting the severity of the accident (no comparative data), the calculated accident risk level is compared with the five-year statistical accident of this section from 2007 to 2011. The comparison results are shown in Fig. 1.

3.3 The Model Validation

In this paper, Spielman correlation coefficient is used to test the correlation between accident number and risk grade. The calculation formula is as follows:

$$\rho = \frac{\sum_{\mathrm{i}} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_i (x_i - \overline{x})^2 \sum_i (y_i - \overline{y})^2}} \tag{10}$$

where

ρ is the correlation coefficient between the number of accidents and the risk level. The greater correlation, the closer the ρ value to ± 1. The calculated correlation

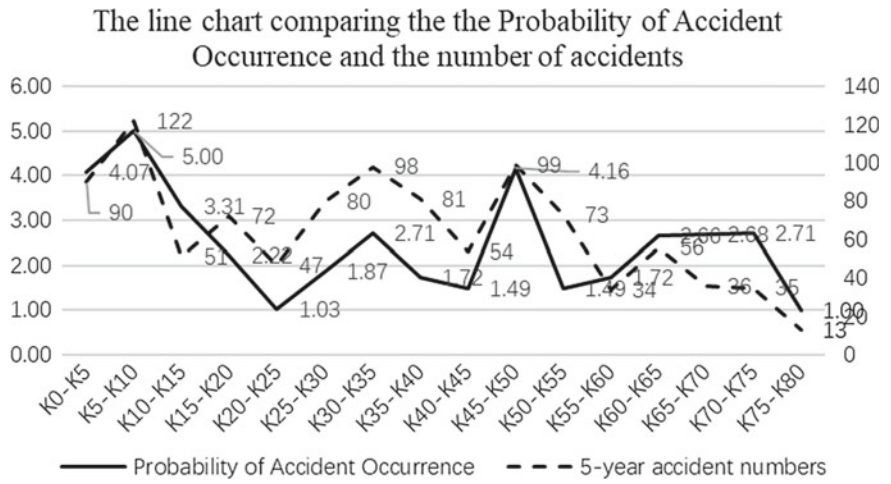


Fig. 1 The line chart comparing the probability of accident occurrence and the number of accidents

reached: $\rho = 0.7109$. Therefore, the evaluation model can be used to evaluate traffic environmental risk.

4 Establishment and Verification of Vehicle Speed Early Warning Model

4.1 Vehicle Speed Early Warning Model

According to the Traffic Environmental Risk Index:

$$TSRI = TSRM \times DFR = E_i \times P_i \times C_i \times \left(\frac{v_i}{v_d}\right)^2$$

Real-time speed can be expressed as:

$$v_i = v_d \times \sqrt{\frac{TSRM}{E_i \times P_i \times C_i}} = v_d \times \sqrt{\frac{TSRM}{\frac{V_i}{V_{\max}} \times \frac{v_{\lim}}{v_{\max}} \times \sum_{j=1}^{4} W_j S_{ij}}} \tag{11}$$

where

$TERM$ Traffic Environmental Risk Model,
E_i Risk Exposure,
P_i Probability of Accident Occurrence,
C_i Severity of Accident,

v_i Real-time Vehicle Speed,
v_d Operating Speed.

Early warning speed considers the following two factors factors:

(1) Early warning speed should be in line with the safe speed most drivers can adapt to, so early warning speed should be as close as possible to the operating speed, and lower than the operating speed.
(2) Early warning speed should be coordinated and stable to avoid large speed difference in different risk sections. Velocity gradient can reflect the coordination of vehicle speed and is closely related to accident rate. Therefore, the speed gradient of early warning vehicle speed should be low.

After the experimental calculation, comparing with 100 sets of data of 16 sections, the warning speed that satisfies the two factors of running speed and speed gradient is the risk index grade of 75 grades. This warning is lower than the average speed of operating speed of 8.36 km/h, the mean velocity gradient value is 1.0482. The speed gradient of Kunshi expressway train is 1.3263. Thus, the speed gradient of the early warning vehicle is lower than that of the operating speed.

The real-time speed corresponding to the traffic environmental risk index of 75 is the early warning speed. Define early warning speed as v. When the Real-time Vehicle Speed v_i is less than early warning speed v ,the vehicle speed system does not give an alarm. When $v_i > v$, The system will give the driver sound and light warning according to the specified risk threshold, that is to say: $\Delta v = v_i - v$.

According to the relevant specifications of road dangerous section identification and the evaluation method of dangerous section, and according to the "road route design specification", vehicle speed warning risk levels are classified I, II, III, IV. The corresponding relationship between risk level and risk threshold is shown in Table 3.

I Speed Early Warning Level, Driver's environmental risk is very low and road environment condition is good. The driver can drive in the road at the desired speed without warning from the system. II Speed Early Warning Level, Drivers are at low environmental risk, and the road conditions are good at this time. Drivers can drive smoothly in the road according to the requirements of speed limit and so on. At this time, the speed warning system only prompts road information such as speed limit. III Speed Early Warning Level, Drivers are in a moderately risky road environment. At this time, the road environment is relatively complex, and drivers need to pay more attention to the driving environment in front of them. Speed warning system

Table 3 The section of road risk level

Risk level	Risk threshold (Δv)
IV (High risk)	$\Delta v > 10$ km/h
III (Moderate risk)	$0 < \Delta v \leq 10$ km/h
II (Mild risk)	-10 km $\leq \Delta v < 0$
I (No risk)	$\Delta v \leq -10$ km/h

can prompt drivers to pay more attention and reduce speed appropriately. IV Speed Early Warning Level, road environmental conditions are dangerous. Vehicle Speed warning system warns drivers that they are in dangerous environment and need to reduce their speed while giving red sound and light alarms.

4.2 Verification of Early Warning Vehicle Speed

The rationality of early warning speed is studied and analyzed from two perspectives: (1) Safety early warning speed should be in line with the range most drivers can bear, so safety early warning speed and operating speed can be compared and analyzed; (2) In the early warning speed, the vehicle should be safer than the operating speed.

According to the calculation model of operating speed, the operating speed of Kunshi expressway (Kunming-Shilin direction) is calculated. The results are shown in Fig. 2.

This picture is based on many experiments and analyses. When the Traffic Environment Risk Index is 75, the corresponding early warning speed is selected. As can be seen from the graph, under the influence of environmental risk, the early warning speed is more gentle than the original running speed, and within the acceptable range of drivers, it is more coordinated to provide reasonable speed for on-road vehicles.

According to the analysis of the spatial distribution characteristics of accident data of Kunshi Expressway in this paper, we can see that K0-K8, K27-K35, K47-K51 are accident-prone sections. According to Fig. 2, the early warning speed of K0-K13 and K48-K50 are obviously lower than the operating speed. This indicates that the Traffic Environment Risk Index in the section is relatively high, which reduces the driver's

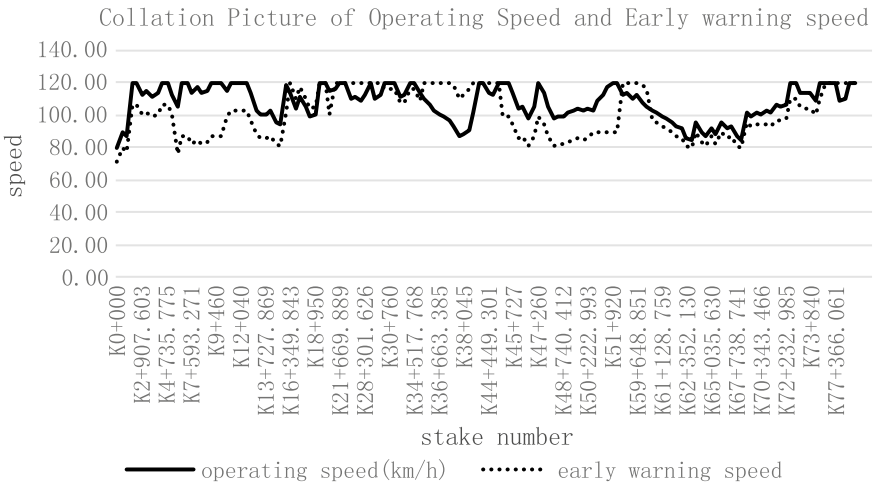


Fig. 2 Collation picture of operating speed and early warning speed

early warning speed accordingly. It is consistent with Kunshi accident-prone section, that is, the speed early warning algorithm can accurately identify the high-frequency accident section, and provide a lower safe speed for drivers.

5 The Design of the Speed Early Warning System

5.1 The Collection and Transmission of the Information of the Speed Early Warning System

With the development of intelligent transportation and automatic vehicles, vehicles have gradually evolved from an independent information island into a node in the intelligent transportation network. The rational use of traffic data information is the main means to reduce traffic accidents and improve the safety of vehicles driving. The speed early warning system is based on high-precision map information, combined with traffic geographic information, network meteorological service information, intelligent traffic perception information, etc., using the traffic environment risk index model for quantitative calculation, through cloud service or on-board system, and the speed early warning is provided for the driver.

① Collection of the road information

This system collects highway road information, and the information mainly include the spatial location information of radius, slope, number of lanes, stadia, bridge and tunnel, traffic buildings and so on. In recent years, the development trend of map navigation is high-precision. Open Drive is a map information file, which is supported by most enterprises. High precision map data based on Open Drive is an important way to acquire road latent risk and a development trend of on-board warning data sources. To realize the collection of road information data, according to the data of Open Drive, the system proposes a method to extract the road information, such as radius, slope, lane change, line-of-sight shielding, bridge and tunnel, safe lane, signal light and toll station, service area, parking area, viewing platform and ramp in the roadside environment. The method is shown in Table 4.

② Collection of traffic facilities data

The road traffic facility information involved in this system mainly includes: guardrail, isolation facility, anti-dazzle facility and line-of-sight guidance facility. Due to the imperfection of map information, there is no suitable technology to directly collect such information data. The risk information involved in traffic facilities mainly identify the design rationality of traffic facilities. Therefore, the collection of traffic facilities data involves image recognition and image processing technology. The system uses computer, image processing, 5G and other high-tech to establish the traffic facilities information database. GPS is used to obtain the location of traffic facilities, and CCD camera is used to collect and save the image information of traffic

Table 4 Information acquisition methods of traffic environmental risk index

Risk factors	Standard layer	Index layer	Partition criterion	Data acquisition path	Data conversion model and principle
Probability of accident occurrence	Road condition	Curve	$R \leq 1000$ m	Open drive	$R = 1/\rho$, where ρ represents the curvature of the road, number of curves with curvature greater than 0.001 in statistical section elements
		Hill	$Slope \geq 3\%$	Open drive	$elev = a + b * ds + c * ds^2 + d * ds^3$, where $elev$ represents the vertical curve of road, a, b, c, d are the parameter of the vertical curve. Let's derive $elev$ once. Gradient with derivative as vertical curve i. when $i \geq 3\%$, see it as a ramp
		Lane change	Increase/decrease in number of lanes	Open drive	Lane Center line is 0, its width is 0.0. lane offset represents lane change
		Insufficient sight distance	Road space sight distance does not meet the requirements of operating speed	Open drive	Object type in open drive marks plants, buildings and other objects. If the outline of the object is within the visual range, there is a risk of visual distance occlusion. The risk level of visual distance occlusion can be calculated by the number of visual distance occlusion in the segment

(continued)

facilities in real time. The image information is saved into the database. At the same time, distance sensor and GPS navigation system information are used to determine the location of the traffic facility information with risks, so as to calculate the risks through the speed early warning system.

Table 4 (continued)

Risk factors	Standard layer	Index layer	Partition criterion	Data acquisition path	Data conversion model and principle
		Bridge/tunnel/culvert	100–300 m before entering tunnel entrance/bridge starting point; Location to 100–300 m after exit of tunnel entrance/bridge terminal	Open Drive	The object types of bridges and tunnels are bridge and tunnel respectively. The names and lengths of bridges and tunnels are recorded in their attributes. Only by counting the number of bridges and tunnels in the section unit, the risk grade of bridges and tunnels in the section can be obtained
		Long downhill without truck escape ramp	Long downhill without truck escape ramp	CDD camera	–
		No signal control intersection	No signal control at intersection	Open drive	The object type is dynamic signal without attribute value yes in junction, which is regarded as non-light-controlled intersection. The number of non-light-controlled intersection in section unit is counted, which is the risk level of non-light-controlled intersection in this section
	Roadside environment	Toll station	The road through the toll station	Open Drive	Toll station, service area, parking area and viewing platform are all building areas in roadside environmental factors. Open Drive marks them as building, specifying outline and name identification, which can be calculated according to building name
		Service areas	The road through the service area		
		Parking area	The road through the parking area		
		Sightseeing stand	The road through the sightseeing stand		
		Expressway ramp	The road through the expressway ramp		

The working principle of image collection of traffic facilities: When a vehicle equipped with the speed early warning system is driving on an expressway, the distance sensor controls the image acquisition card and the color camera to capture a scene image of the road ahead, accepts the distance sensor and GPS, and saves the transportation facility information into the database. These all are used to provide transportation facility information for the speed early warning system. The information collection system of transportation infrastructural is shown in Fig. 3.

③ Collection of weather information

In this paper, autonomous navigation technology is used to collect weather information in rainy, snowy and foggy days, it mainly includes road tracking, obstacle detection and positioning, etc. The involved sensors mainly include CDD infrared sensor, infrared sensor, laser 3d imaging radar, millimeter wave radar and so on. In order to meet the need of all-weather information acquisition, the system uses infrared sensors to perceive and identify the weather information. Summarize, the collection of the weather information is for visibility, air temperature, surface temperature, humidity, rainfall, which requires a a multiple sensor fusion system. And this system choose the distributed fusion method, and establish the information acquisition system of weather, its structure as shown in Fig. 4.

④ Acquisition of traffic flow information

Traffic flow information includes flow and speed information. They are acquired from the navigation system, which can collect accurate location information, speed,

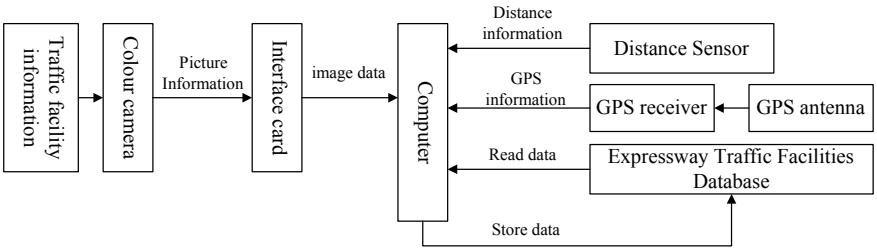


Fig. 3 Information collection system of transportation infrastructural

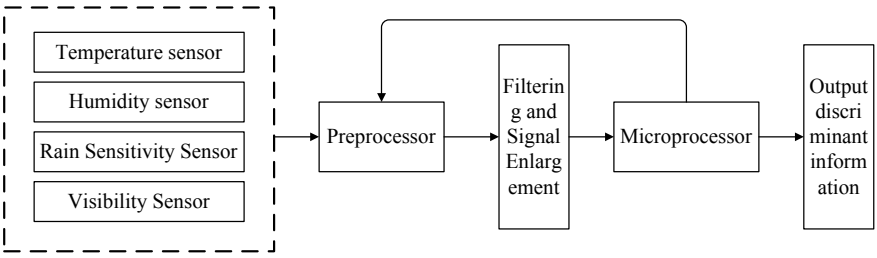


Fig. 4 Information acquisition system of weather

time and other traffic information. In this paper, the speed early warning system takes the vehicles on the highway as the carrier, and the speed early warning system client as the data acquisition terminal. Traffic flow information acquisition uses the communication between clients for data transmission, so as to achieve the maximum range of data acquisition and analysis under the condition of low cost. With the help of wireless network data communication, the speed early warning system transmits the vehicle data o the background server. After the server judges the user ID, the corresponding database is stored. In the database, the vehicle ID number is used as the index to store the traffic flow data of the vehicle. The traffic flow data needs GPS navigation map matching, and different vehicle position points are classified into different sections, so as to analyze the real-time traffic flow in the section and the maximum traffic flow in the region.

Information transmission of early warning system:

(1) 5G Cellular Network Technology

5GCellular network refers to the fifth generation mobile communication network. Its peak theoretical transmission speed can reach 10 Gb per second. It can meet the requirements of vehicle communication and vehicle-road coordination. Therefore, 5G technology will be widely used in the future development of vehicle network. Vehicle networking will become an important application area of 5G network. The wireless network designed by the speed early warning system in this paper uses 5G cellular network technology to ensure the accuracy and real-time of environment and vehicle information. The basic principle is to use 5G cellular network technology to achieve high reliability, low delay and active interaction between vehicle and road information. It mainly realizes the transmission of traffic data acquired by GPS navigation system in vehicle speed early warning system. By using 5G cellular network technology, the message transmission between vehicle and cloud terminal can be realized, and the network fusion and information fusion can be realized in a real sense.

(2) CAN

Because the speed early warning system in this paper is based on traffic environment risk, it has the characteristics of large amount of information and complex data format. The information exchange between vehicle speed early warning system and peripheral equipment and electronic control unit is extremely complex. In order to solve the problem of information transmission and improve the accuracy of information, the controller local area network CAN (Controller Area Network) is adopted in this system. Controller Local Area Network (CAN) is a communication technology designed by Bosch Company for information exchange among various electronic control devices in automobiles. The direct communication distance of AN can reach 10 km/Skbps and the maximum speed can reach 1 Mbps/40 m. PHILIPS is the first to introduce a new generation of SJAl000 with more complete functions. There are two working modes: Basic CAN mode (82C200 compatible mode) and PriCAN mode (extended feature), which can complete all functions of CAN bus physical layer and data connection layer.

③ TMS320C54x Digital Signal Processor

The main function of the vehicle speed early warning system in this paper is to make traffic environment risk analysis and early warning display for on-road vehicles. It needs a lot of road traffic information as the data base. After data preprocessing and analysis, the analysis results are warned on the page of the vehicle early warning system. The function is accomplished jointly by the background database and the operation and analysis center.

This system uses the digital signal processor of TMS320C54x to process and analyze the traffic environment information. Its structure adopts improved Harvard structure, CPU with dedicated hardware logic, on-chip memory, on-chip peripherals, and a high performance instruction set. C54x has the following main features: 6-channel DMA controller, 3 multi-channel buffer serial ports (McBSPs), user-configurable 8_bit or 16_bit main port interface (HPl8/16), and 128 kxl6_bit in-chip RAM.

5.2 Workflow of Vehicle Speed Early Warning System

Vehicle speed early warning system integrates GPS positioning system, information transmission network layer, information acquisition application layer, on-board computer and acousto-optic alarm system to perceive road traffic environment risk and vehicle running status such as its own speed. Through the calculation of risk prediction model, real-time risk perception of on-the-road vehicles is carried out and reasonable early warning of vehicle speed or speed is given. Through the assistant function of vehicle speed early warning system, driver's risk perception of road environment can be improved, and the possibility of traffic accidents can be reduced. Safety of vehicles on the road has been improved. The structure of the speed warning system is shown in Fig. 5.

Real-time vehicle speed corresponding to different risk levels can be obtained from risk level threshold. When the risk level is high, the real-time vehicle speed of the model is relatively low. On the contrary, when the risk level is low, the real-time vehicle speed is relatively high.

The principle of speed warning based on traffic environmental risk index is shown in Fig. 6.

6 Conclusion

(1) The traffic environmental risk index model was established, which can evaluate the environmental risk in the road. The traffic environment risk index includes three aspects: the risk exposure, the probability of accident occurrence and the severity of accident. And the dynamic risk factors are used to modify the model, including real-time driving speed and running speed.

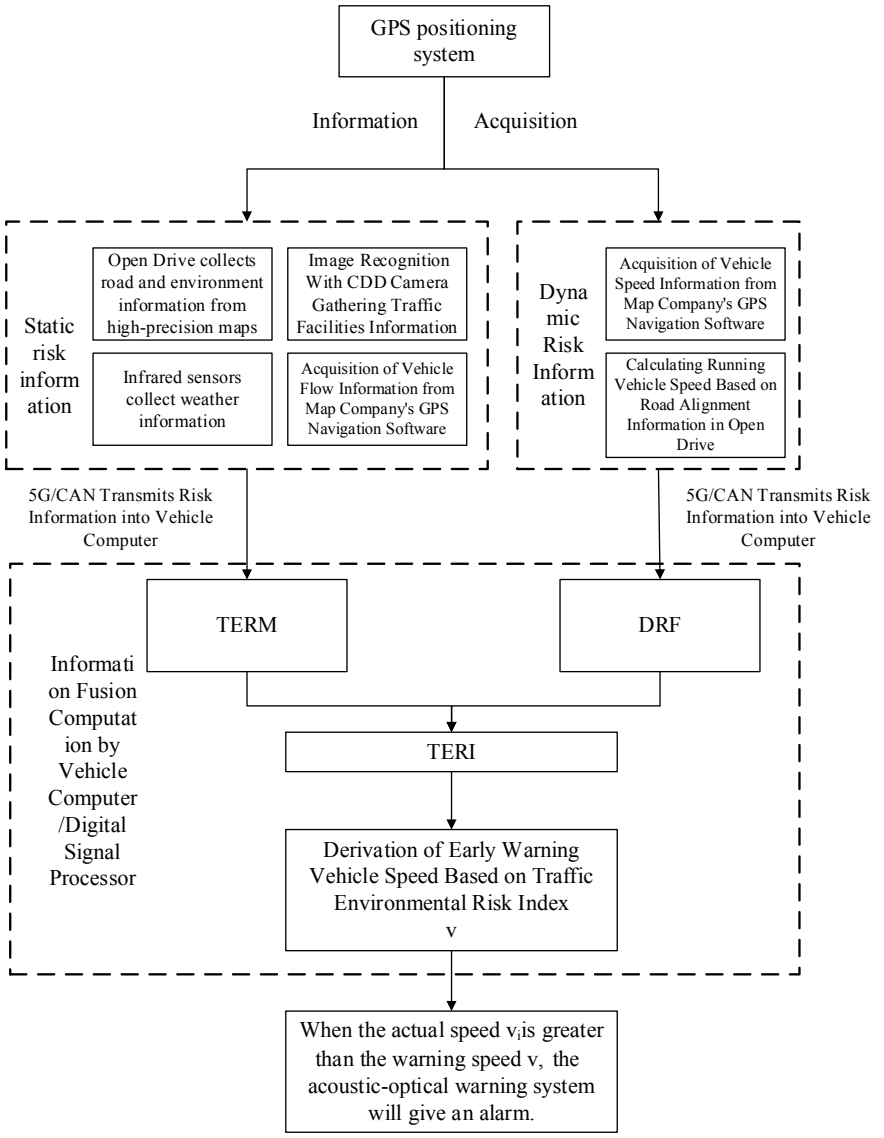


Fig. 5 Framework of vehicle speed early warning system

(2) The validity and feasibility of the traffic environment risk index model were tested and verified. Through obtaining and verifying the environmental risk data and accident data of Kunshi Expressway, it is found that the prediction result of the risk level of Kunshi Expressway by this model has a correlation of 0.7109 with its accident data. It is proved that this model can effectively predict the risk status of vehicles.

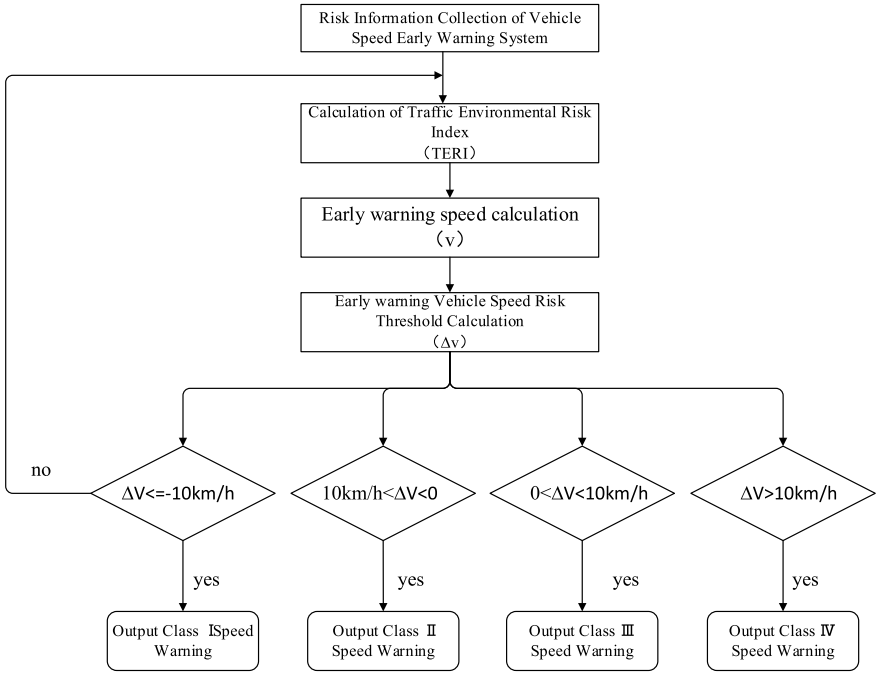


Fig. 6 The theory of vehicle speed early warning

(3) The speed early warning model was established and verified, and the speed early warning system based on the speed early warning model was designed. The speed early warning model is based on the traffic environment risk index, and combined with the evaluation indexes of running speed and speed gradient, etc. Through Open Drive high-precision map, CDD camera, GPS navigation map, infrared sensor, 5G cellular network and other technologies, the speed early warning system finishes the collection and transmission of road information, traffic facilities, weather conditions, traffic flow and other information. The on-board computer and the digital signal processor are used to analyze the risk information and calculate the warning speed, and the speed early warning is provided for the driver according to threshold of the warning speed.

References

1. Zhao L, Jia W, Dai S et al (2018) Characteristics of urban road traffic safety in China. Urban Transp China 16(3)
2. Cui FY, Xin SX, Feng XC (2015) Study on comprehensive evaluation theory of vehicle internal environment for traffic safety and design of vehicle early warning system. High-tech Commun 25(2):171–179

3. Shen B, Zhang Z, Liu H et al (2018) Research on a conflict early warning system based on the active safety concept. J Adv Transp 2018:1–11
4. Huang XL, Liu YL, Duan J et al (2017) Road traffic safety risk assessment based on grey correlation and fuzzy comprehensive evaluation method. Math Pract Cogn 47(7):208–215
5. Wang JJ, Cao XD, Yang YF (2018) Risk isolation analysis of mountain highways areas based on CRSG model. China J Highw Transp 31(09):119–128
6. Yang CF, Gao HN, Sun JS et al (2016) Traffic safety evaluation method for expressways based on improved grey correlation-combination weighting method. J Expressway Transp Res Dev (English Edition) 10(4):71–77
7. Xiong J, Guo FX, Wan HS et al (2017) A test and evaluation method for driver's latent risk perception ability CN106491144A
8. Jung S, Qin X, Noyce D (2010) Rainfall effect on single-vehicle crash severities using polychotomous response models. Accid Anal Prev 42:213–224
9. Chen Feng, Chen Suren (2010) Reliability-based assessment of vehicle safety in adverse driving conditions. Transp Res Part C 19(1):156–168
10. Leur PD (2002) Development of a road safety risk index. Transp Res Rec 1784(1):33–42

The Prediction of Delay Time Class Caused by CTCS-3 Onboard System Fault Based on Decision Tree

Lijuan Shi, Ang Li and Liquan Chen

Abstract The faults of train control system will lead to delay, which will affect the operational efficiency of the railway network. In this paper, the decision tree algorithm (CART) is used to predict the delay time level caused by CTCS-3 Onboard System Fault, which takes the location of train failure, the fault component of CTCS-3 on-board system, the fault phenomenon of CTCS-3 on-board system as data features. In the natural language fault record, based on expert experience, extract the key features needed and grade the delay time. The selected features are put into the decision tree algorithm for classification and prediction, SMOTE algorithm is used to solve the problem of unbalanced number of categories, and grid search algorithm is used to adjust the model parameters. Finally, the output results of the algorithm are analyzed. The decision tree model yields a classification accuracy of 76% for the given data of fault feature and can be considered for delay time level prediction caused by CTCS-3 system fault. From the experimental results, the proposed method can be recommended for the prediction of the delay time level caused by CTCS-3 system fault.

Keywords CTCS-3 on-board system fault · Delay time class · Expert knowledge · Decision tree

L. Shi (✉)
Shanghai Key Laboratory of Rail Infrastructure Durability and System Safety, Tongji University, No. 4800 Caoan Road, Jiading District, Shanghai 201804, People's Republic of China
e-mail: shilijuan150@163.com

A. Li
Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, No. 4800 Caoan Road, Jiading District, Shanghai 201804, People's Republic of China
e-mail: Lyon123456@126.com

L. Chen
China Railway Guangzhou Group Co., Ltd., Guangzhou, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_85

1 Introduction

CTCS3 on-board equipment is the core part of CTCS3. There are some fault types and faulty equipment. Studying the regularity of the faulty events is helpful to formulate the emergency plan after the event. Considering that CTCS3-300T is widely used in China's high-speed railway, this paper takes CTCS3-300T on-board equipment as an example to study.

In the field of fault prediction, decision tree is widely used. Lee et al. [1] developed a supervised decision tree method which uses the machine learning and data mining techniques. It is designed to estimate the key factors in knock-on delays. Once the delay root cause and the interactions among the factors responsible for these are classified, some reactions can be taken to enhance the punctuality of railway operations such as timetable adjustment. Guo [2] summarize train's fault data and fault diagnosis methods. After the fault component and constraints over data sigma completeness can be converted to attributes of conditions and decision making in the decision tree (C4.5), then the regulations between fault features and sigma completeness can be found. After extract effective features from real-time data, we can detect the similarity for fault features, so the faults would be located. Xiao et al. [3] apply the decision tree to realize the train's rail deformation detection. The attribute of train's rail deformation detection is selected and the track condition is divided into two categories: track fault and good condition. It belongs to decision tree dichotomy. Madhusudana [4] presents the classification of healthy and faulty conditions of the face milling tool using Decision tree (J48 algorithm) technique based on machine learning approach. Rabah et al. [5] proposed a method using decision tree algorithm to detect and diagnose the faults in grid connected photovoltaic system. Three numerical attributes and two targets are chosen to form the final used data, the attributes are temperature ambient, irradiation and power ratio, the first target is either healthy or faulty state, the second contains four classes labels named free fault, string fault, short circuit fault or line-line fault for diagnosis. Jiang et al. [6] use an adaptive local root mean square (RMS) calculation method which has been used to extract the time-domain characteristics of diesel engine cylinder head vibration signal. The classification model of the diesel engine based on the classification and regression tree (CART) algorithm was established.

Due to the limited literature on the application of decision tree in train control system to judge faults and delays, and the lack of research on the prediction of delay time, it is necessary to find the relationship between faults and delay time.

2 Technology Roadmap

Delay time grade prediction model is constructed in three stages, as shown in Fig. 1, expert knowledge, machine learning and other related knowledge are used. In Phase 1, expert knowledge is applied to extract data features from 192 natural language

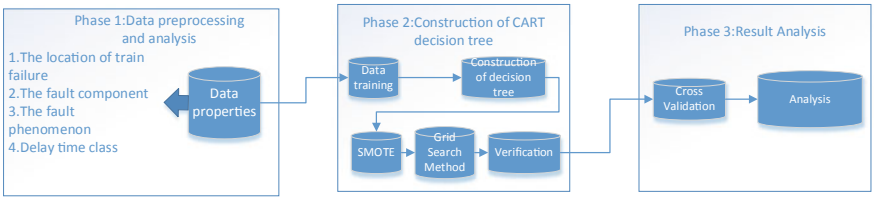


Fig. 1 Technology roadmap

records of CTCS3-300T fault event. In Phase 2, CART algorithm is used to construct decision tree. SMOTE algorithm is used to solve the problem of sample category imbalance in data. The grid search algorithm finds the optimal parameters of the model. In Phase 3, the data features are substituted into the model and the output results are analyzed.

3 Fault Feature Extraction Based on Expert Knowledge

There are many faults in on-board equipment, but not all of them are caused by on-board equipment itself. There are many other factors leading to abnormal operation of equipment, so it is particularly important to determine the fault component and fault phenomenon. This study is based on 192 CTCS-300T fault records. According to the content of "Basic Requirement for Emergency Management of Running Failure of EMU of Shanghai Railway Administration", the organization and analysis of D-Series High-Speed Trains whose failure time exceeds 20 min should be carried out. Therefore, in this study, the delay time in fault data is divided into 2 levels: delay time longer than 20 min and delay time no more than 20 min. From the data, 162 data with delay time no more than 20 min and 30 data with delay time longer than 20 min are available.

3.1 Data Set Features

In 192 existing CTCS3-300T fault data, the main fault status is as follows:

(1) Software defects of on-board equipment have not been thoroughly solved.
In order to solve the problem of software defects occurring from time to time in train operation the manufacturers are trying their best to carry our software upgrade and optimization work, and the effect is remarkable. Some software problems have been basically solved, but there are still software faults such as inconsistent A/B codes.

(2) There are many Balise Transmission Module (BTM) faults

BTM receives the ground transponder information and sends it to ATPCU and C2CU for processing. BTM type faults occur in a large number, besides, the types and phenomena of faults are also complicated. Common BTM failures include invalid BTM ports, BSA errors, BTM test timeouts, etc.

(3) There are many faults in train interface equipment

On-board train interface equipment includes Vital Digital Input & Output (VDX), Multifunction Vehicle Bus (MVB), relay and so on. VDX is the train interface, which is used to input and output safety-related signals such as output emergency braking and collecting feedback of braking when the train exceeds speed. VDX has redundant configuration, VDX1 and VDX2. There are two main faults: VDX message invalidation and VDX port invalidation.

3.2 Fault Feature Extraction

Through the analysis of the existing fault data, combined with the composition of the equipment and expert knowledge, the fault locations and fault phenomena in the data are shown in Table 1. In this study, the data feature will be extracted according to Table 1.

Fault occurring at a station or between stations also affects the length of the delay. Therefore, in addition to the fault component and fault phenomenon, the data feature also include the fault location. The fault location is divided into two types: at a station and between stations.

4 Constructing Decision Tree Model by CART

Decision tree is a case-based inductive learning algorithm, which aims at inferring classification rules represented by decision tree from a set of disordered and irregular cases. The purpose of constructing decision tree is to find out the relationship between attributes and categories, and to use it to predict the categories of future data. CART algorithm is one of the classical algorithms used to construct decision tree.

4.1 CART Algorithm

Decision tree (CART algorithm) uses Gini index to select partition attributes. The purity of data set D can be measured by Gini value.

Table 1 Fault information of CTCS3-300T on-board system

Fault component	Fault reason	Fault phenomena
Balise Transmission Module (BTM)	1. TM hardware failure 2. The BTM port is invalid 3. BSA 4. Balise lost	1. Host and DMI communication is interrupted 2. Balise message error, brake parking 3. CTCS2 failure 4. Starting failure 5. Brake parking
Compact Antenna Unit (CAU)	1. CAU antenna failure	1. Host and DMI communication is interrupted 2. Brake parking 3. Starting failure 4. CTCS2 failure
Viral Computer (VC)	1. ATPCU software failure 2. ATPCU hardware failure 3. C2CU software failure	1. Starting failure 2. Brake parking 3. Host and DMI communication is interrupted 4. CTCS2 failure, brake parking
Track Circuit Reader (TCR)	1. Poor TCR transmission 2. TCR status is abnormal 3. TCR information is unreasonable 4.TCR hardware failure	1. CTCS2 software failure,brake parking 2. Host and DMI communication is interrupted 3. TCR failure
Driver-machine Interface (DMI)	1. DMI hardware failure 2. DMI communication fault 3. DMI software failure	1. Host and DMI communication is interrupted 2. Brake parking 3. DMI starting failure
Speed and Distance unit (SDU)	1. Radar fault 2. Speed sensor fault 3. SDU hardware failure 4. ODO no service	1. Speed sensor fault 2. Radar fault 3. Brake parking 4. Host and DMI communication is interrupted
Train interface equipment: 1. Vital Digital Input & Output (VDX) 2. Relay 3. Multifunction Vehicle Bus (MVB)	1. VDX message is invalid 2. VDX port is invalid 3. Relay failure 4. MVB	1. Host and DMI communication is interrupted 2. Brake test failed 3. Brake parking 4. Starting failure
Juridical Recorder Unit (JRU)	1. JRU hardware failure 2. JRU software failure	1. Brake parking 2. JRU failure
Software	1. Inconsistent A/B codes	1. Host and DMI communication is interrupted 2. Brake parking

$$\begin{aligned} \text{Gini(D)} &= \sum_{k=1}^{|y|} \sum_{k' \neq k} p_k p_{k'} \\ &= 1 - \sum_{k=1}^{|y|} p_k^2 \end{aligned} \tag{1}$$

p_k—The proportion of class k samples in the current sample set D is $p_k (k = 1, 2, \ldots, |y|)$.

Intuitively, Gini(D) reflects the probability that two samples are randomly extracted from data set D with different class labels. Therefore, the smaller the Gini(D), the higher the purity of the data set D.

The Gini index of attribute a is defined as:

$$\text{Gini}_{\text{index(D,a)}} = \sum_{v=1}^{V} \frac{|D^v|}{|X|} Gini\left(D^v\right) \tag{2}$$

Therefore, in the set of candidate attributes A, the attributes that minimize the Gini index after partitioning are selected as the optimal partitioning attributes. That is

$$a_* = argmin\ Gini_{index(D,a)} (a \in A) \tag{3}$$

After the Decision tree is constructed, the data volume of the two types of data with a delay time longer than 20 min and no more than 20 min are quite different. Considering the class imbalance of the data set, the SMOTE algorithm is used to synthesize some new sample for the class with less quantity.

4.2 SMOTE Algorithm

The basic idea of the SMOTE algorithm is to analyze the data in the category with small number of samples and synthesize new samples based on the small number of samples. The algorithm flow is as follows:

(1) For each sample X in the category with few samples, calculate the distance to all samples of the same category by Euclidean distance, and obtain its K-Nearest Neighbors.
(2) Set a sampling ratio according to the sample imbalance ratio to determine the sampling magnification N. For each minority sample X, randomly select several samples from its k-nearest neighbors, assuming that the selected neighbor is X_n.
(3) For each randomly selected neighbor X_n, new samples is constructed with the original sample according to the Formula (4).

$$x_{new} = x + rand(0, 1) \times (\tilde{x} - x) \tag{4}$$

In this experiment, the sampling ratio is set to 0.8, that is, the data amount of the category with a small number of samples accounts for 80% of the data amount of larger category after the data is generated by the SMOTE algorithm. Solved the problem of sample category imbalance, and used the generated sample to construct the decision tree model (CART algorithm).

4.3 Adjust Parameters Using Grid Search Method

Since the parameters of the decision tree will affect the final result, the grid search algorithm is used to adjust the parameters of the model. The grid search method is a basic parameter optimization algorithm. It essentially divides the parameters to be searched into a grid with the same length in a certain coordinate range according to the proposed coordinate system. Each point in the coordinate system represents a set of parameters, which can be verified by substituting each point in the given interval into the decision tree model to derive parameters that optimize the performance of the overall system.

According to the grid search algorithm, the maximum depth of the decision tree is set to 8, the minimum number of leaf nodes is set to 1, and the criterion is the Gini index. Based on these parameters, the best accuracy is obtained.

5 Verification and Results Analysis

Usually, we will evaluate the generalization error of the learning machine through experiments, and we need to use the test set for this purpose. In order to improve the generalization ability, “cross validation” is adopted. The data set is divided into k mutually similar subsets, each time using the union of k-1 subsets as the training set, and the remaining subset as the test set, so that it can be performed k times of training and testing, the final return is the mean of the k test results.

According to the above decision tree construction process, a decision tree for delay time level prediction can be obtained in Fig. 2.

5.1 Results Analysis

Table 2 shows the detailed accuracy of the CART model. The precision is based on the prediction results, that is, how many data are predicted to be correct in a certain type of data. The recall rate is based on the sample, which represents the proportion of the correct predictions in a certain type of sample. Overall, the classification accuracy

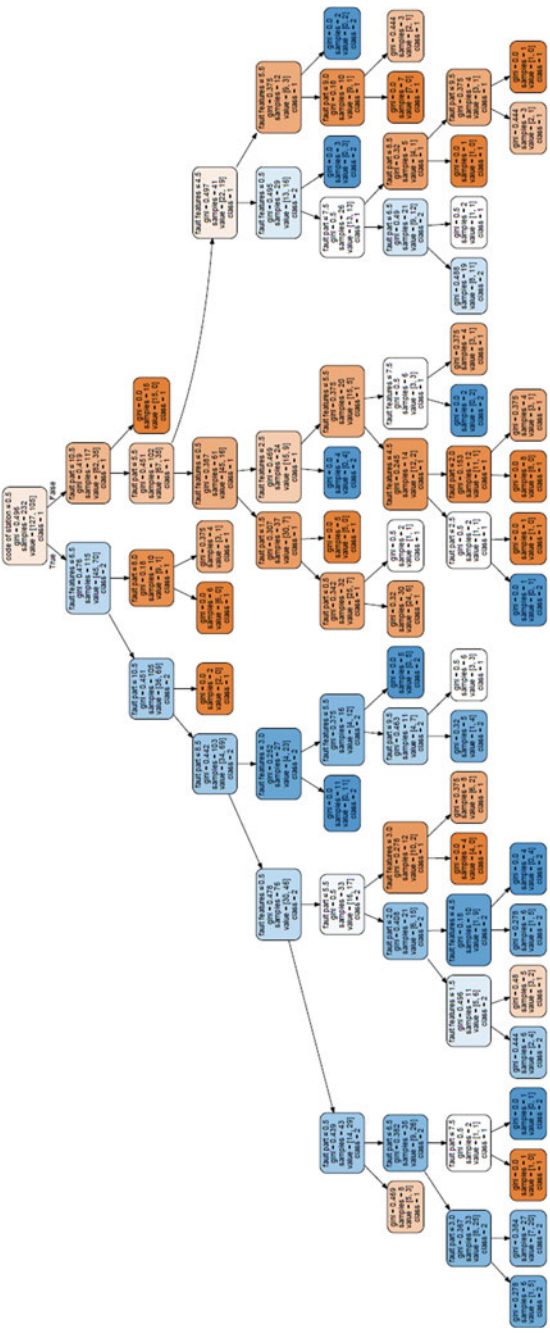


Fig. 2 Decision tree

Table 2 Detailed accuracy classification of decision tree

	Precision	Recall	f1-score
Less than 20 min	0.72	0.75	0.73
More than 20 min	0.82	0.80	0.81

of the decision tree model reached 76%, that is, in 232 data (the training set included new data generated), a total of 176 data were correctly classified. This model can be considered for delay time level prediction analysis.

6 Conclusion

In this paper, through the machine learning method, using decision tree and expert knowledge, the prediction of delay time grade caused by CTCS3-300T Onboard System Fault is studied. The correspondence between fault components and fault phenomena is summarized, and the data feature are extracted, which contains fault components, fault phenomena and the location fault occur. The decision tree is constructed, and the Gini index is used as the classification basis. According to the data structure, the SMOTE algorithm is used to generate the data, and the grid search algorithm is used to calculate the optimal model parameters. The experimental results show that the decision tree has a good classification accuracy of 76%, thus, the model can be recognized in the field of delay time level prediction. Therefore, the decision tree method can be used to predict the delay time level caused by CTCS3-300T on-board equipment fault.

Acknowledgements Authors would like to acknowledge the support of the research program of ***Comprehensive Support Technology for Railway Network Operation (2018YFB1201403)***, which is a subproject of ***Advanced Railway Transportation Special Project*** belonging to the ***13th Five-Year National Key Research and Development Plan*** funded by Ministry of Science and Technology of China.

References

1. Lee WH, Yen LH, Chou CM (2016) A delay root cause discovery and timetable adjustment model for enhancing the punctuality of railway services. Transp Res Part C: Emerging Technol 73:49–64
2. Guo W (2017) Research on train fault early warning method based on data Sigma completeness of train control system
3. Xiao Q, Wang C, Liang H et al (2006) Application of C4.5 algorithm in train's rail deformation detection. Comput Technol Develop 16(4)
4. Madhusudana CK, Kumar H, Narendranath S (2018) Fault diagnosis of face milling tool using decision tree and sound signal. Mater Today Proc 5(5):12035–12044
5. Rabah B, Samir M (2018) Fault detection and diagnosis based on C4.5 decision tree algorithm for grid connected PV system. Solar Energy 173:610–634

6. Jiang Z, Wei D, Wang L et al (2018) Fault diagnosis of diesel engines based on a classification and regression tree (CART) decision tree. J Beijing Univ Chem Technol (Natural Science) 45(4):71–75

Road Network Equilibrium Analysis Based on Section Importance and Gini Coefficient

Fei Su, Xiaofang Zou, Yong Qin, Shaoyi She and Hang Su

Abstract Traffic flow on road network has a character of disequilibrium under the road network structure, traffic flow distribution and so on. In order to quantify its imbalance, taking section importance as the index, the model for road network equilibrium analysis is proposed based on the Gini coefficient. First, considering the road network structure, traffic flow distribution and the influence between sections, the section importance measurement based on space-time influence and space-time distribution is constructed to reflect the critical level of sections in the road network. Second, the road network equilibrium is discussed through Gini coefficient and Lorenz curve. Finally, the proposed model is applied in a subset of Beijing's road network, and the results show that the model is simple and flexible to evaluate road network equilibrium in different dimensions. It has great significance for mastering the distribution law of traffic flow, optimizing road network structure, adjusting traffic capacity allocation and improving the efficiency of road network resources.

Keywords Traffic flow · Road network equilibrium · Space-time influence · Space-time distribution · Section importance

F. Su · S. She · H. Su
China Transport Telecommunications & Information Center, Beijing 100011, People's Republic of China

National Engineering Laboratory of Transportation Safety & Emergency Informatics, Beijing 100011, People's Republic of China

F. Su · Y. Qin
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, People's Republic of China

X. Zou (✉)
China Merchants New Intelligence Technology Co., Ltd, Beijing 100073, People's Republic of China
e-mail: zouxiaofangfei@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_86

1 Introduction

Traffic congestion has been heavily fixed eyes on, especially in Beijing, Shanghai and other cities. Traffic congestion usually occurs in some certain areas or sections. The spatial location and capacity of these certain sections often have a great impact on the traffic state of the surroundings. Traffic flow on road network has an obvious character of disequilibrium, and it can be treated as one of the reasons for leading to road traffic congestion. Therefore, road network equilibrium analysis is an important work in the research on traffic congestion. It can provide useful support for mastering the distribution law of traffic flow, optimizing road network structure, adjusting traffic capacity allocation and improving the efficiency of road network resources. It has important theoretical significance to alleviate traffic congestion.

Many researchers have paid attention to the analysis of road network equilibrium. Dai et al. [1] presented an approach to calculate the unbalance of flow distribution considering freeway network topology, travel route choice behavior and traffic flow. Sun et al. [2] investigated the time and space distribution characteristics of the traffic congestion and bottlenecks in different network topologies (e.g., small world, random and regular network). Deng et al. [3] shown that the analysis of highway network structure characteristics based on the complex network theory can reflect route importance from the view of connectivity, centrality, intermediate and reliability. Shen [4] used complex network theory to measure the disequilibrium of a road network structure by node degree, betweenness, and tightness. He et al. [5] analyzed the unbalance of highway network structure relying on the index of highway network area density and transport density.

In general, there has been an increase in the application of road network equilibrium analysis. It is turned out that the methodologies work very well. However, most of them just analyze the problem of equilibrium from the point of view of network structure, traffic volume or microcosmic, without considering the effects of other factors, for example, the influence of traffic flow between sections. It leads to an inaccurate outcome somewhat, because traffic flow can be regarded as a macroscopic phenomenon emerging under the action of many factors.

For this, and with the goal of quantifying the traffic flow imbalance of road network, the model for equilibrium analysis is proposed by taking section importance as the index. It can be considered that the more similar the section importance is with others, the more balanced the road network will be. The article is structured as follows: the importance of traffic flow equilibrium analysis of road network is emphasized in Sect. 1. Section 2 gives the measurement of section importance. For a section, its importance is measured by two aspects: space-time influence and space-time distribution. Space-time influence is to explain its influence on other surrounding sections, and space-time distribution is to explain its contribution to the whole network. Section 3 gives a mathematical description and the modeling process of the equilibrium analysis model. In Sect. 4, the model has been tested in practice on a subset of Beijing expressway section and the results are analyzed in detail. Finally, we make some conclusions with discussions on future directions in Sect. 5.

2 Section Importance

In this paper, the section importance is measured by two aspects based on the spatial and temporal features of traffic flow: space-time influence and space-time distribution. Space-time influence is measured to compute the strength of influence between section and its neighbors. Space-time distribution can be explained as how much the traffic state of a section contributes to the whole road network.

2.1 Space-Time Influence

In 2008, Box et al. [6] employed the cross-correlation function to measure the correlation of two spatial objects at a given delayed time. As a correlation measure, the CCF treats two traffic time series as a bivariate stochastic process and measures correlation of traffic flow between a certain section in the road network and its neighbors at s time lags. Given two time series X and Y, the CCF at s time lags can be denoted as follows:

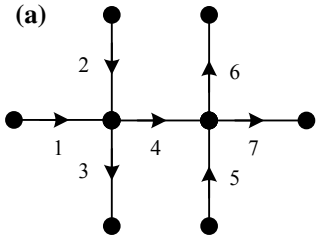
$$\gamma(s) = \frac{E(x(t) - \mu_X)(y(t+s) - \mu_Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}} \tag{1}$$

where μ_X, μ_Y, σ_X^2, σ_Y^2 are the means and variances of time series X and Y respectively. $x(t)$ is the observations of X at time t, and $y(t+s)$ is the observations of Y at time $t+s$.

However, there are usually not only one kth-order neighbor of a certain section in the network. For this, the spatial weight matrix is drawn into the CCF, to represent the space-time influence between a section and its kth-order neighbors clearly and accurately.

Spatial weight matrix is one of the mathematical models to measure the adjacency between any two spatial locations in the network, such as network topological, adjacency relation and so on [7]. Drawing from graph theory, a network can be viewed as a graph $G = (N, E)$, where N is a set of n nodes, E is a set of edges connecting pairs of nodes [8]. In fact, road network is a complex network composed of sections and intersections, where the intersection represents connection between the sections. Therefore, in this paper, the adjacency matrix can be expressed in the way as below: i and j are edges with variable observations and the nodes represent connections between them. When i and j are directly linked by a node, it can be called first-order neighbors, and the adjacency matrix containing all first-order relations between spatial locations of a network is termed its first-order adjacency matrix W_1. Second-order adjacency matrix W_2 of a network is the first-order matrix of its first-order adjacency matrix and so on. By following ways, adjacency matrix W_k can be defined.

Besides, to define the incidence structure, the direction of influence should be also considered. In our opinion, road network can be viewed as a directed graph when



(b)

Segment	1	2	3	4	5	6	7
1	0	0	1	1	0	0	0
2	0	0	1	1	0	0	0
3	1	1	0	0	0	0	0
4	1	1	0	0	0	1	1
5	0	0	0	0	0	1	1
6	0	0	0	1	1	0	0
7	0	0	0	1	1	0	0

(c)

Segment	1	2	3	4	5	6	7
1	0	1/2	0	0	0	1/2	1/2
2	1/2	0	0	0	0	1/2	1/2
3	0	0	0	1	0	0	0
4	0	0	1/2	0	1/2	0	0
5	0	0	0	1	0	0	0
6	1/2	1/2	0	0	0	0	1/2
7	1/2	1/2	0	0	0	1/2	0

Fig. 1 Spatial weight matrices of road network: **a** first-order spatial weight matrix; **b** second-order spatial weight matrix

considering the traffic flow [9]. Furthermore, the adjacency matrix of road network is different from other directed networks:

(1) The traffic only flows from upstream to downstream, but the influence may be occurred in both directions. In free conditions, the influence will be from upstream to downstream, while in congested conditions, it will be mainly from downstream.
(2) When two sections flowing into or out of the same link from the same direction, the influence may be mostly from the second-order instead of first-order neighbors.

In this work, the kth-order adjacency matrix weights are denoted as Eq. (2), considering the distance between two sections, to explain the road network structure. Taking the road network shown in Fig. 1 as an example, the spatial weight matrix can be obtained as in Fig. 1 through the definition of adjacency matrix and spatial weights.

$$w_{ij} = \begin{cases} \frac{1}{k}, \text{ when } i \text{ and } j \text{ are } k\text{th-order neighbors} \\ 0, \text{ else} \end{cases} \tag{2}$$

On the basis of it, given the traffic flow time series X of a section and the time series Y weighted its kth-order spatial neighbors at s time lags, the CCF can be developed as follows.

$$\gamma_i(k,s) = \frac{\sum_{t=1}^{T}\left(x_i(t)-\overline{x_i}\right)\left(W^{(k)}x_i(t+s)-\overline{W^{(k)}x_i}\right)}{\sqrt{\sum_t\left(x_i(t)-\overline{x_i}\right)^2}\sqrt{\sum_t\left(W^{(k)}x_i(t+s)-\overline{W^{(k)}x_i}\right)^2}} \tag{3}$$

where $\gamma_i(k,s)$ is the space-time correlation between section i and its kth-order neighbors at time lag s; T is the period of statistical time; $x_i(t)$ is traffic flow observations of the given section i at time t; $\overline{x_i}$ is the average value of time series X; $W^{(k)}x_i(t+s)$ is a space delay operator, representing the weighted mean traffic observations of all kth-order neighbors of the section i at time lag s, and $\overline{W^{(k)}x_i}$ is the average value of

the kth-order neighbors' traffic, then it can be denoted as follows:

$$W^{(0)}x_i(t) = x_i(t),\ W^{(k)}x_i(t) = \sum_{j=1}^{m} w_{ij}^{k} x_j(t) \tag{4}$$

where m is the number of the section in the network, w_{ij}^{k} is the weight between section i and j in the kth-order spatial weight matrix.

Note that space-time correlation measures the influence of one section on its kth-order neighbors under a time lag s, and it can be treated as a discrete process varying with time delay s and space delay k. In order to calculate the influence of a section on other sections around in the network comprehensively, it is necessary to synthesize space-time influences in spatial and temporal dimensions respectively. Therefore, an integrated spatial weight matrix is considered in spatial dimension. It can be obtained by the sum of different order spatial weight matrix as the following equation:

$$W' = \sum_{i=1}^{k} W_i \tag{5}$$

where W_i is the ith-order spatial weight matrix. W' is the integrated spatial weight matrix, representing that all indirect or direct influence of one section on its neighbors from first to kth-order will be considered simultaneously. It is worth noting that the weights w'_{ij} should be row standardized in the case study.

In addition, the technique for order preference by similarity to an ideal solution (TOPSIS) is employed as a synthetical method for space-time correlation in temporal dimension to obtain the space-time influence. Its steps are as follows [10].

(1) Defining positive and negative ideal points

$$\begin{aligned} A^{+} &= \left\{ \max_i r_i(s) | s \in \mathbf{S}, 1 \le i \le N \right\} = \left\{ A^{+}(s) | s \in \mathbf{S} \right\}, \\ A^{-} &= \left\{ \min_i r_i(s) | s \in \mathbf{S}, 1 \le i \le N \right\} = \left\{ A^{-}(s) | s \in \mathbf{S} \right\} \end{aligned} \tag{6}$$

where $r_i(s)$ is space-time correlation of section i at time lag s synthesized in spatial dimension. S is the set of time delay values. N is the number of sections in the road network. $A^{+}(s)$ is the maximum value of space-time influences. $A^{-}(s)$ is the minimum value of space-time influences. In fact, it is usually carried out that $A^{+}(s) = 1$, $A^{-}(s) = -1$, based on the range of $[-1, 1]$ of space-time correlation.

(2) Computing distance

The Euclidean weighted distance is usually used as follows to calculate the distance between space-time correlation and the positive and negative ideal points.

$$E_i^+ = \sqrt{\sum_s w_s(r_i(s) - A^+(s))^2}, 1 \le i \le N, \quad E_i^- = \sqrt{\sum_s w_s(r_i(s) - A^-(s))^2}, 1 \le i \le N \tag{7}$$

where w_s is weighting coefficient. E_i^+ is the distance between space-time correlation and the positive ideal point. E_i^- is the distance between space-time correlation and the negative ideal point.

(3) Calculating the influence degree

The relative closeness e_i of space-time correlation of section i to the ideal point is calculated to measure the influence degree of section i on other sections in the road network.

$$e_i = \frac{E_i^-}{E_i^- + E_i^+}, 1 \le i \le N \tag{8}$$

(4) Normalization of the influence degree

So as to distinguish the differences of the influence of a section on others easily, the influence degree should be normalized as follows.

$$c_i = \frac{e_i - \min e_i}{\max e_i - \min e_i}, \ 1 \le i \le N \tag{9}$$

where c_i is the space-time influence of the section i.

2.2 Space-Time Distribution

For example, at a certain moment, the heavier the traffic congestion of a section (that is, the greater the traffic volume, or the longer the travel time, or the lower the travel speed of the section), the greater the contribution of the section to the road network traffic state, which also reflects the greater importance of the section in the road network [11]. Meanwhile, traffic flow changes periodically, and it changes with time in each cycle. Therefore, the traffic flow observation at every moment in a cycle should be considered, when computing the space-time distribution. Given the traffic flow observations of a road network, the space-time distribution can be denoted as follows:

$$d_i = \frac{s_i - \min s_i}{\max s_i - \min s_i}, 1 \le i \le N \tag{10}$$

where d_i is the space-time distribution of section i. s_i expresses the contribution of the section i to the road network in a time period T. $\max s_i$ is the maximum value

of all sections' contributions to the road network. mins_i is the minimum value of all sections' contributions to the road network. N is the number of sections in the road network.

$$s_i = \sum_{t=1}^{T} \frac{\theta_x^T(i,t)x_i(t)}{w_x^N(t)}, \theta_x^T(i,t) = \frac{x_i(t)}{\sum_{t=1}^{T} x_i(t)}, w_x^N(t) = \sum_{i=1}^{N} x_i(t) \tag{11}$$

where $x_i(t)$ is the traffic flow observation of section i a certain time t. $\theta_x^T(i,t)$ is the traffic flow temporal distribution of section i a certain time t in the statistical time period. $w_x^N(t)$ is the sum of real-time traffic flow observations in all sections of road network at a certain time t. T is the statistical time period.

2.3 Importance Calculation

In this work, the section importance is measured by two aspects in this work: the contribution of a section to the whole network and its influence on other surrounding sections. Thus, space-time influence and space-time distribution should be considered at the same time to calculate the importance of the section. In order to get the comprehensive value for the importance of a section in the road network, the linear weighting method is employed, and it can be denoted as follows:

$$I_i = \alpha \cdot c_i + (1-\alpha) \cdot d_i, 1 \le i \le N \tag{12}$$

where I_i is the importance measurement of the section i. α is the weight, determining the key factor in measuring the section importance. When the weight α is large ($\alpha > 0.5$), it means that more attention will be paid to space-time influence of the section on its surroundings; When the weight α is small ($\alpha < 0.5$), it means that more consideration will be paid to space-time contribution of the section to the road network.

3 Road Network Equilibrium Analysis

Gini coefficient was proposed by Italian economist Corrado Gini in 1921 on the basis of Lorentz curve [12]. It can be used to measure the equilibrium degree of resource allocation such as property, capital, product, market and so on. As shown in Fig. 2, taking accumulated percentage of component as abscissa, and accumulated percentage of section importance as ordinate, Lorentz curve forms a square with an area of 1. The diagonal line is the absolute average line. The Gini coefficient is the proportion of the area surrounded by Lorentz curve and absolute average line to the

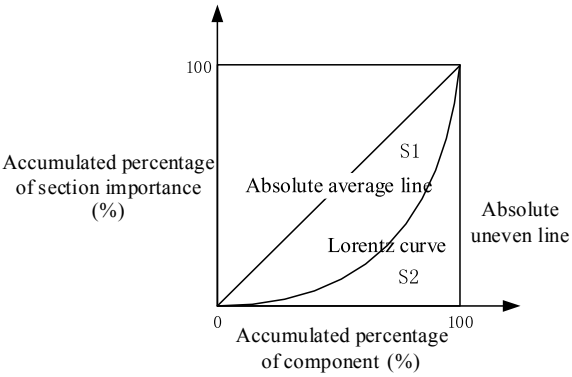


Fig. 2 The grading standards of Gini coefficient

area between absolute fairness line and absolute unfairness line. The greater Gini coefficient is, the more unbalanced the road network is.

The method of road network equilibrium analysis based on Gini coefficient can be described as follows.

(1) According to the factors such as section, road or statistical period etc., the road network X is divided into n groups $X_i = (\alpha_i, \beta_i)$, $i = 1, 2, \ldots, n$ with an ascending sort order, where α_i is the number of components in X_i, β_i is the standardized section importance of X_i.

(2) Calculating the proportion of components A_i and accumulated percentage A'_i.

$$\mathrm{A}_i = \alpha_i \Big/ \sum_{i=1}^{n} \alpha_i, \mathrm{A}'_i = \sum_{j=1}^{i} \mathrm{A}_i \tag{13}$$

(3) Calculating the proportion of section importance B_i and accumulated percentage B'_i.

$$\mathrm{B}_i = \beta_i \Big/ \sum_{i=1}^{n} \beta_i, \mathrm{B}'_i = \sum_{j=1}^{i} \mathrm{B}_i \tag{14}$$

(4) Taking A'_i as the abscissa and B'_i as the ordinate, the expression of Lorentz curve $\mathrm{B}' = f(\mathrm{A}')$ is obtained by fitting scatter points $\left(\mathrm{A}'_i, \mathrm{B}'_i\right)$.

(5) Computing Gini coefficient.

$$G = S_1/(S_1 + S_2) = S_1/0.5 = 1 - 2\int_0^1 \mathrm{B}'\mathrm{dA}' \tag{15}$$

(6) Computing the slope θ_i between two adjacent groups X_i. It means that the bigger θ_i is, the quicker the change of X_i is and the more unbalanced it is.

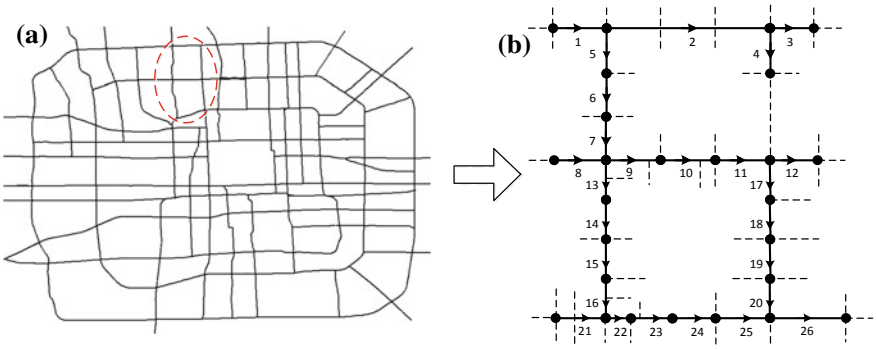


Fig. 3 A subset of Beijing's road network as the test network: **a** the location of the test network, **b** the section in the test network

$$\theta_i = \left(\mathrm{B}'_i - \mathrm{B}'_{i-1}\right) \Big/ \left(\mathrm{A}'_i - \mathrm{A}'_{i-1}\right) = \mathrm{B}_i \Big/ \mathrm{A}_i \tag{16}$$

4 Case Study

4.1 Experiment Scenario and Data Set

The model for road network equilibrium analysis is applied to traffic data in practice on Beijing's road network. The Expressway Traffic Information Detection System built by Beijing Traffic Management Bureau collects real-time volume, speed and time occupancy for expressways depending on the microwave detectors. These sensors cover nearly all the expressways including ring roads and connecting expressways. A subset of Beijing's road network is chosen as test network in the case study. As mentioned in reference to Fig. 3 and Table 1, the test network comprises 26 sections.

Traffic flow data for six weeks from 9 May to 19 June 2011 is selected in practice. For the survey data, it is obtained daily from the detectors every 2 min. In order to reduce the random elements caused by the noise in the 2-min data, the traffic data is transformed into 10 min discrete time intervals. That is, there are 144 observations in one day.

4.2 Network Equilibrium Analysis

In order to analyze the equilibrium of the road network in detail, the modeling is performed according to the four periods of a day: 0:00–24:00, AM peak (7:00–10:00), Interpeak (11:00–14:00) and PM peak (17:00–20:00). It has been widely accepted

Table 1 The origin and destination of the 26 sections in the test network

Section	Origin	Destination	Section	Origin	Destination
1	Zhan Chun Second Bridge	Xue Yuan Bridge	14	Ming Guang Bridge North	Ming Guang Bridge
2	Xue Yuan Bridge	Jian Xiang Bridge	15	Ming Guang Bridge	Wen Hui Bridge
3	Jian Xiang Bridge	Bei Chen Bridge West	16	Wen Hui Bridge	Xi Zhi Men Bridge
4	Jian Xiang Bridge	Jian Xiang Bridge North	17	Ma Dian Bridge	Bei Jiao Market
5	Xue Yuan Bridge	Hua Yuan North Road	18	Bei Jiao Market	Xin Kang Road
6	Hua Yuan North Road	Xue Zhi Bridge	19	Xin Kang Road	An De Road
7	Xue Zhi Bridge	Ji Men Bridge	20	An De Road	De Sheng Men Bridge
8	Capital University of Physical Education and Sports	Ji Men Bridge	21	Bei Zhan Bridge	Xi Zhi Men Bridge
9	Ji Men Bridge	Hua Yuan Road	22	Xi Zhi Men Bridge	Huapi Factory Hutong
10	Hua Yuan Road	Bei Tai Ping Zhuang Bridge	23	Huapi Factory Hutong	Xin Jie Kou
11	Bei Tai Ping Zhuang Bridge	Ma Dian Bridge	24	Xin Jie Kou	Ji Shui Tan Bridge
12	Ma Dian Bridge	Yu Min Middle Road	25	Ji Shui Tan Bridge	De Sheng Men Bridge
13	Ji Men Bridge	Ming Guang Bridge North	26	De Sheng Men Bridge	Drum Tower Bridge

in transport studies that traffic behaves differently in each time period. The analysis is threefold: Firstly, the importance of 26 sections in the test network for 42 days is calculated separately; Secondly, the imbalance of the test network is quantified by using section importance; Finally, road network equilibrium is analyzed according to different statistical times.

(1) Section importance calculating

Due to the length of the sections, the neighbors of a certain section higher than five orders may be outside one-time lag (it means 10 min in this case study). Therefore, in this case study, five spatial weight matrices (W_1–W_5) are employed to obtain the integrated spatial weight matrix W' according to Eq. (5). Note that the weights should be row standardized in the case study.

On the basis of Chap. 2, space-time influence and space-time distribution of the 26 sections for the 42 days are computed in the case study. Taking 20 May 2011 as an example, Fig. 4 gives the space-time influences of the 26 sections performed according to 0:00–24:00, AM peak, Interpeak and PM peak, respectively. As shown in Fig. 5, space-time distribution corresponding to the 26 sections on 20 May 2011 is also given as an example.

Section importance is measured both by space-time influence and space-time distribution. According to the linear weighting method mentioned above, Table 2 lists the measurements of the section importance for each section during the 4 periods of a day on 20 May 2011 (Here the weight $\alpha = 0.5$).

(2) Network imbalance quantification

As shown in Fig. 3, the test network can be divided into 26 groups. That is, each group contains one section. Therefore, taking the 26 groups as an object, the imbalance of the test network for the 42 days is quantified based on Gini coefficient and Lorenz curve. The modeling is performed according to 0:00–24:00, AM peak (7:00–10:00), Interpeak (11:00–14:00) and PM peak (17:00–20:00) respectively. Taking 20 May 2011 as an example, the section importance of the 26 groups is given in

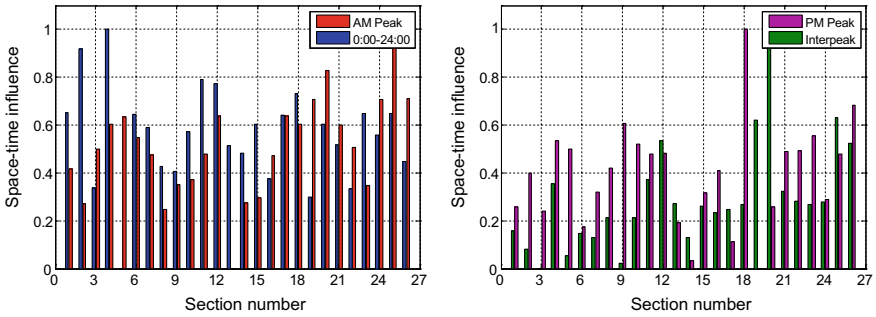


Fig. 4 Space-time influence of the 26 sections on 20 May 2011

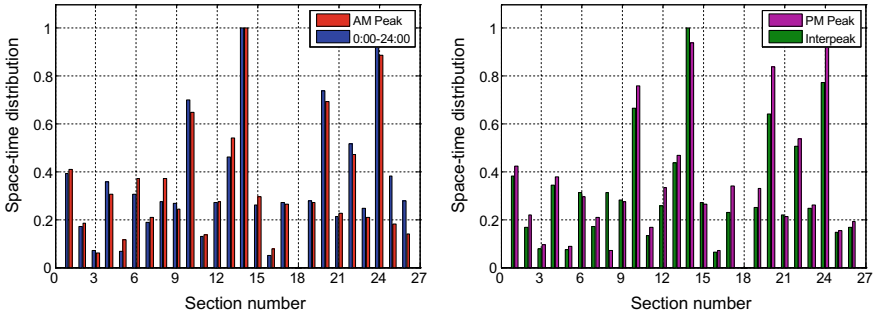


Fig. 5 Space-time distribution of the 26 sections on 20 May 2011

Table 2 The results of section importance on 20 May 2011 when $\alpha = 0.5$

Section (Group) X_i	Section Number α_i	Section importance β_i			
		0:00–24:00	AM Peak	Interpeak	PM Peak
1	1	0.5223	0.4141	0.2701	0.3417
2	1	0.5447	0.2284	0.1253	0.3092
3	1	0.2031	0.2814	0.0397	0.1666
4	1	0.6791	0.4551	0.3496	0.4574
5	1	0.0344	0.3744	0.0645	0.2935
6	1	0.4745	0.4601	0.2300	0.2355
7	1	0.3888	0.3428	0.1515	0.2648
8	1	0.3514	0.3101	0.2637	0.2463
9	1	0.3360	0.2962	0.1523	0.4414
10	1	0.6373	0.5102	0.4396	0.6399
11	1	0.4600	0.3082	0.2535	0.3244
12	1	0.5216	0.4564	0.3964	0.4085
13	1	0.4884	0.2699	0.3531	0.3304
14	1	0.7407	0.6371	0.5653	0.4862
15	1	0.4320	0.2950	0.2672	0.2900
16	1	0.2126	0.2754	0.1483	0.2418
17	1	0.4554	0.4512	0.2387	0.2267
18	1	0.3661	0.3016	0.1335	0.5000
19	1	0.2889	0.4902	0.4356	0.1646
20	1	0.6703	0.7616	0.8204	0.5476
21	1	0.3652	0.4138	0.2716	0.3501
22	1	0.4264	0.4904	0.3942	0.5162
23	1	0.4475	0.2784	0.2572	0.4078
24	1	0.7542	0.7975	0.5260	0.6452
25	1	0.5144	0.5905	0.3907	0.3165
26	1	0.3636	0.4246	0.3459	0.4365

Table 2. According to the method of road network equilibrium analysis based on Gini coefficient, the 26 groups are sorted in ascending order by section importance β_i corresponding to the 4 periods separately. Then, the proportion and cumulative percentage of section number and section importance for each period are calculated. Table 3 gives the measurements of Lorentz curve and Gini coefficient for the 4 periods. Figure 6 shows the fitting process of the Lorentz curve and the slope θ of each group (section) for the 4 periods. Examining these figures and tables, the following statements could be drawn:

- Table 3 shows a comparison of Gini coefficient for the 4 periods on 20 May 2011. It could be observed that the value of Gini coefficient in the Interpeak is the largest.

Table 3 The Lorentz curve and Gini coefficient for the 4 periods on 20 May 2011

	Lorentz curve	Gini coefficient
0:00–24:00	$0.51x^2 + 0.5x - 0.02$	0.203
AM Peak	$0.53x^2 + 0.43x + 0.015$	0.186
Interpeak	$0.78x^2 + 0.19x - 0.00012$	0.296
PM Peak	$0.57x^2 + 0.41x + 0.0062$	0.199

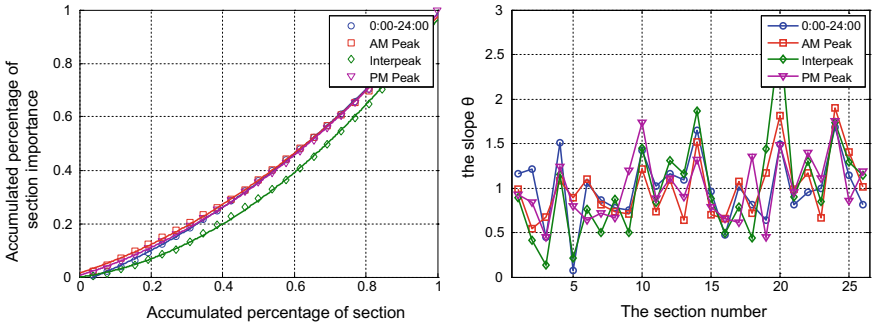


Fig. 6 The Lorentz curve and the slope for the 4 periods on 20 May 2011

That is, compared with other periods, the test network is the most unbalanced during the Interpeak on 20 May 2011.

- Figure 6 displays the slope of each group (section) for the 4 periods the change trends of slope are similar to that of section importance. It is clear that the slope of section 20 is the largest during the Interpeak, and it is 2.71. It means that the importance of section 20 is the greatest, and the change of importance of section 20 is the quickest.
- It could be regarded that section 20 may be one of the main factors of leading to the road network imbalance. In other words, it needs to be paid more attention to in daily management.

(3) Network equilibrium analysis

Taking the 26 groups as an object, the Gini coefficient of the road network for the 42 days is also computed in this case study, corresponding to the four periods of a day: 0:00–24:00, AM peak (7:00–10:00), Interpeak (11:00–14:00) and PM peak (17:00–20:00). In order to analyze the equilibrium of the road network in detail, the average Gini coefficient is list in Table 4 based on the different statistical periods. Compared with the for four periods, the average Gini coefficient during the Interpeak is the largest basically. That is to say, the test network is the most unbalanced corresponding to the Interpeak. It may be caused by the nonstationary traffic flow in the Interpeak. On one hand, travel route is selected randomly based on the people's daily travel habits when traffic is relatively smooth. The distribution of traffic flow in road network is stochastic and unbalanced. On the other hand, there are strong differences in the

Table 4 The average Gini coefficient of the different statistical periods

	The 42 days	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
0:00–24:00	0.212	0.207	0.208	0.202	0.218	0.221	0.209	0.221
AM Peak	0.236	0.250	0.242	0.226	0.224	0.250	0.241	0.217
Interpeak	0.268	0.249	0.291	0.259	0.270	0.263	0.261	0.286
PM Peak	0.242	0.247	0.241	0.253	0.242	0.242	0.257	0.216

Table 5 The highest frequency section with the largest slope in the different statistical periods

	The 42 days	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
0:00–24:00	24	24	24	24	24	24	24	24
AM Peak	24	24	20	24	24	10/24	14	14
Interpeak	20	20	20	14/20	20	20/24	24	24
PM Peak	24	24	24	24	20	24	14/24	24

traffic state of sections during the Interpeak, and its distribution is also unbalanced. The strength of influence between sections is obviously different.

In addition, the sections with the largest slope is counted in this case study. Table 5 gives the highest frequency section with the largest slope according to different statistical periods. It could be observed that the highest frequency section with the largest slope mainly occurs on sections 24, 20 and 14 no matter which periods. It means that these sections are the bottleneck of test road network. In general, the daily traffic of these sections is larger, and they are more likely to affect the traffic state of the test road network.

5 Conclusion

Road network has an obvious character of disequilibrium under the network structure, traffic flow distribution and so on. It can be treated as one of the reasons for leading to road traffic congestion, because the congestion usually occurs in some certain areas or sections, which often has a great impact on the surrounding sections. Therefore, road network equilibrium analysis is an important work in the research on traffic congestion. It can provide effective decision support for enhancing the supervision over the important sections, and carrying out reasonable measures for optimizing the road network structure. Besides, it can help traffic management to advance measures of improving the efficiency of road network resources to alleviate traffic congestion.

The goal of this work is to analyze the equilibrium of road network. Taking section importance as the index, the model is proposed based on the Gini coefficient and Lorentz curve, and that goal is clearly achieved by this effort. The advantages of the framework are as follows:

(1) The measurement of section importance is proposed. It is determined by the contribution of section itself to the network and the influence of the section others.
(2) Taking section importance as the index, the network equilibrium analysis not only considers the spatial differences of sections in the road network, but also considers differences of traffic flow distribution and influence.

Moreover, the method applied to in subset of Beijing road network shows that it is simple and flexible to evaluate road network equilibrium in different time and space dimensions. It can also be easily applied to other frequency data or experiment scenarios, because they share the common theories foundation. All these features make this approach appealing and with plenty of potential for improving. The next steps of this work are as follows:

(1) Continue to analyze and discuss the reasons for the imbalance of road network further.
(2) To improve the method by considering multiple factors, for example, OD traffic, section capacity and so on.
(3) To apply this work to the traffic state evaluation, traffic guidance, traffic organization and other research. For example, summarizing adaptive adjustment model of the weight in traffic state evaluation according to the imbalance of road network.

Acknowledgements This work was supported in part by the National Key Research and Development Program of China (Grant No. 2017YFC0803900).

Declaration of Conflicting Interests
The authors declare that there is no conflict of interest regarding the publication of this article.

References

1. Dai HN, Yao EJ, Liu SS et al (2017) Flow inequality of freeway network based on Gini-coefficient. J Transp Syst Eng Inf Technol 17(1):205–211
2. Sun HJ, Wu JJ, Ma D et al (2014) Spatial distribution complexities of traffic congestion and bottlenecks in different network topologies. Appl Math Model 38:496–505
3. Deng YJ, Yang YF, Ma RG (2010) Highway network structure characteristics based on complex network theory. China J Highway Transp 23(1):98–104
4. Shen HF (2012) An approach to analysis and evaluation of highway network structural properties for risk assessment and emergency management. Beijing Jiaotong University, Beijing
5. He YH, Fan BQ, Dong JS et al (2010) Highway network distribution equilibrium based on Gini coefficient. J Transp Syst Eng Inf Technol 10(6):163–168
6. Box G, Jenkins G (1976) Time series analysis: forecasting and control. Wiley, San Francisco, USA, pp 256–268
7. Wang YF, He HL (2007) Spatial data analysis method. Science Press, Beijing, pp 56–57
8. Peeters D, Thomas I (2009) Network autocorrelation. Geogr Anal 41(4):436–443

9. Kamarianakis Y, Prastacos P (2005) Space-time modeling of traffic flow. Comput Geosci 31(2):119–133
10. Wu CF (2015) Research on the evaluation of urban road traffic state based on the combination method. Jiangxi University of Science and Technology, Ganzhou, pp 26–30
11. Petter H (2003) Congestion and centrality in traffic flow on complex networks. Adv Complex Syst 6(2):163–176
12. He BQ, Hong XJ (2016) A review of Gini coefficient calculation and its decomposition method. Stat Decis 14:13–17

Research on Driving Workload Characteristics of Drivers Under Various Dangerous Scenarios Based on EEG

Shumin Feng and Bin Sheng

Abstract In order to study the driving workload characteristics of drivers under various dangerous scenarios, the electroencephalography (EEG) data of drivers under different dangerous scenarios are analyzed based on real driving experiments. The ErgoLAB human-machine environment synchronization platform was used to collect and extract the driver's EEG signal. Meanwhile the appropriate EEG index was selected. The statistical analysis method was used to study the EEG data of the drivers, and the EEG variation rates of the driver under various dangerous scenarios was obtained to reflect driving workload. Two conclusions are reached from the experiment designed to figure out the relationship between driving workload and dangerous scenarios. Two conclusions were reached from analyzing indicators. First, various dangerous scenarios have significant impacts on driving workload. When there exists one dangerous scenario with fewer non-motor vehicle protection measures easily causing serious traffic accidents such as pedestrians or bicycles, there will be a higher driving workload. Besides, the impact on the EEG variation rates of the driver in various dangerous scenarios caused by the external factors such as turning and encountering pedestrian (bicycle) or a relatively sound protection measures taken like vehicle interactions is small, resulting in lower driving workload. Second, driving workload is not only affected by dangerous scenarios but also by driving experience and the age of driver.

Keywords Dangerous scenarios · Driver · EEG · Driving workload

1 Introduction

There are four main factors affecting traffic accidents: drivers; vehicles; roads; environmental conditions. Driver-related factors account for more than 90% of total traffic accidents. In order to reduce traffic accidents caused by driver's dangerous

S. Feng (✉) · B. Sheng
School of Transportation Science and Engineering,
Harbin Institute of Technology, Harbin, China
e-mail: zlyfsm2000@sina.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_87

behaviors, studying the driving workload is of great importance. According to relevant paper, EEG as an intuitive response to the human's mental status is adopted repeatedly. Aghajani H. and Omurtag A. used electroencephalography (EEG) and functional near-infrared spectroscopy (FNIRS) to assess mental load, and the results showed that the classification accuracy of the Hybrid system include the EEG and FNIRS signal was greater quantify mental load [1]. Andrei G. L. and Takashi S. from Nissan et al. used the event-related potential P3 latency index to evaluate the drivers' mental load during human-computer interaction, and found that the P3 latency index was not affected by external load [2]. Lim W. L. et al. designed and implemented an experiment for mental workload recognition to collect EEG data, they proposed a method based on support vector machine for combining statistics and fractal dimension (FD) features, which has the best monitoring accuracy of mental load [3]. Some who studied the relationship of EEG and drivers' driving behaviors and workload had marvelous accomplishments. Kim Il-Hwa et al. studied the detection of braking intentions in different emergency situations based on the combination of EEG features in a simulated driving environment [4]. S. Sega et al. proposed a multiple linear regression equation for driving workload respect to vehicle-related variables [5]. Kim J. Y. et al. considered the influence of the vibration characteristics of the vehicle on brain waves, and studied the driving load changes under different road types on the basis of removing the interference waves [6]. Kim H. S., Hwang Y., Yoon D., et al. studied driving workload differences of five kinds of drivers' behavior, including left-turn section, right-turn section, rapid-acceleration section, rapid-deceleration section, and lane-change section [7]. Besides, some scholars have studied the relationship between driving workload and other physiological parameters except EEG. Yuan Wei studied the impact of urban road types on driving workload [8]. Wu Meng carried out a research on the factors which affected drivers' driving workload in expressway ramp area [9]. Hu J. and Wang R. studied the influence mechanism of different weather conditions on driving workload based on the driving simulator [10]. Wang Chang et al. studied that under the specific circumstance where pedestrians suddenly invaded the road, with fixed speed of the pedestrian, the drivers' psychological load increased due to short pedestrian intrusion distance [11].

Currently, most of the research on driving workload uses driving simulator for analysis, and mainly focuses on indicators such as vision, electrocardiogram, respiration, and skin electricity. The research on EEG and involved in dangerous scenarios is relatively rare. This paper adopts the natural driving experiment method to collect the EEG signals of drivers under various dangerous scenarios. The driving workload is analyzed based on the EEG variation rates of the driver. The purpose of this paper is providing theoretical basis for driving assistance system and reducing driving workload.

Table 1 Basic information of the participants

Band	Number	Age (year)		Driving age (year)	
		Mean	Standard deviation	Mean	Standard deviation
Total	14	36.93	12.91	7.64	7.74
Male	10	35.3	13.59	7.7	8.86
Female	4	41	11.74	7.5	4.93

2 Experimental Design and Data Acquisition

2.1 Experimental Preparation

In order to obtain experimental data, drivers with different ages and driving ages were invited to take part in the experiment. During natural driving under urban road environment, collect the changes in the brain electrical characteristics of the drivers under various dangerous scenarios.

A total of 14 participants, 10 males and 4 females were recruited in the experiment. All of them had obtained motor vehicle driving licenses, and their vision or corrected vision was normal. All the drivers in the experiment had certain driving experience but with different driving proficiency. Besides, they can complete the driving task independently and safely. The age of drivers ranged from 24 to 59 years (mean = 36.93, standard deviation = 12.91), and the driving age was between 1 and 27 years (mean = 7.64, standard deviation = 7.74) (Table 1).

2.2 Experimental Route

There are a large number of motor vehicles on urban roads, where non-motor vehicles and pedestrians often appear. And the traffic environment is complex, which is prone to various dangerous scenarios. Therefore, the natural driving of each participant was selected on the urban road, and the free flow of traffic was guaranteed as possible as we can when designing the experimental route. After cautiously consideration, the experimental route chosen is the main road in Nangang District of Harbin, which is a bi direction and six lane road.

2.3 Experimental Steps

The testers who wore electroencephalograph and Tobii Pro Glass2 eye tracker drove cars equipped with high-definition camera (collecting data of driving operation and external conditions) and vehicle information collection sensor (collecting data of

Fig. 1 Natural driving experiment scene

vehicle speed, acceleration, etc.). Before the experiment started, the testers were briefly introduced the purpose and general flow of the experiment. Besides, the testers was worn electroencephalograph and debugged. There was a 5–10 min vehicle familiar process on campus to assist testers in adapting as soon as possible, in order to eliminate tension and fear. Finally, the testers drove vehicle into the experimental route. Because the experimental route was a city road, the traffic environment was complex and continuous driving was required. For the safety of the testers, the test time is 1.25 h, and the traveling speed was limited to 40 km/h. During the experiment, the following staffs recorded the dangerous scenarios information including time and place, without affecting the driver's normal driving operation.

The instruments worn and natural driving experiment scene are shown in Fig. 1.

2.4 Selection of Dangerous Scenarios

Due to the different participants of the urban traffic environment, the dangerous scenarios are generally divided into four types: motor vehicle-motor vehicle, non-motor vehicle-vehicle, motor vehicle-pedestrian, and motor vehicle-external environmental. The frequency of dangerous scenarios was analyzed and high frequency of them were extracted. After extracting video, combined with the classification of 37 types of pre-crash dangerous scenes proposed by NHTSA (National Highway Traffic Safety Administration), the classification results (Table 2) found that the top 5 accounts for about 89% of the dangerous scenarios.

Table 2 Types and proportions of dangerous scenarios

Types of dangerous scenarios	Number	Proportion (%)
Front car off lane	77	24.2
Vehicle changing lane	67	21.1
Pedestrian (bicycle) conflict with prior vehicle maneuver	55	17.3
Front car decelerating suddenly	45	14.2
Pedestrian (bicycle) conflict without prior vehicle maneuver	39	12.2
Others	35	11.0
Total	318	100

The explanations of dangerous scenarios are as follows:

(1) Front car off lane. The main behavior is that the car is traveling straight, and the preceding car driving in the same lane deviates from the lane and invades the road of the vehicle.
(2) Vehicle changing lane. The main behavior is the lane change of the car, which encroaches on the roads of other vehicles traveling in the same direction.
(3) Pedestrian (bicycle) conflict with prior vehicle maneuver. The main behavior is the encounter of pedestrians (bicycles) during the turn of the vehicle.
(4) Front car decelerating suddenly. The main behavior is that the car follows the preceding car in a straight line, and the front car suddenly decelerates.
(5) Pedestrian (bicycle) conflict without prior vehicle maneuver. The main behavior is that the vehicle is traveling in a straight line and encounters a pedestrian (bicycle).

3 EEG Index Extraction and Processing

3.1 EEG Indicators

EEG measures voltage fluctuations caused by the flow of ionic currents in brain neurons. EEG signals can reflect the electrical activity of the brain tissue and the functional status of the brain. It is mainly composed of α waves, β waves, θ waves, δ waves, including basic characteristics such as frequency, interval, amplitude and phase. The characteristics of the four typical EEG signals are shown in Table 3.

The α wave is characterized by stable and calm mood, high level of consciousness activity with strong work ability; the β wave represents sensitivity to the external environment, emotional, anxiety, alter, concentrated to the external environment, and engaged in high intelligence activities; the θ wave shows consciousness, and the body is deep and relaxed; the δ wave characterizes deep sleep, unconscious state.

Table 3 Characteristics and location of typical EEG signals

Band	Frequency (Hz)	Amplitude (μV)	Characteristics	Location
Delta (δ)	0.2–3.99	20–200	Sleep	
Theta (θ)	4–7.99	20–100	Drowsiness	Frontal, temporal
Alpha (α)	8–12.99	10–100	Relaxed	Occipital, parietal
Beta (β)	13–30	5–25	Excited, busy	Frontal, parietal, temporal

3.2 EEG Indicators Processing

The sample frequency of the EEG is 250 Hz. During the collection of EEG signals, there are interferences from other physiological signals. For example, electrooculogram (EOG) exists in the entire collection of EEG signals, so it is necessary to eliminate noise components. The EDF file of the EEG signal was derived from MATLAB and EEGLAB plug-in, which obtained the amplitude values of the multiple electrode positions. First, the noise reduction processing was performed, and then the EEG signal was converted into the time domain to the frequency domain by the fast Fourier transform. In the frequency domain range, the range was extracted according to the range to which each band belongs, and finally the power spectral density mean of each waveform was calculated. According to previous studies, Alpha waves and Beta waves can fluctuate significantly in dangerous scenarios. Therefore, Alpha waves/Beta waves were selected as indicators of driving workload to analyze EEG.

The formulas for calculating the average power spectral density of each frequency band (taking Alpha waves as an example) are as follows:

$$G(\alpha) = \int_{f_d}^{f_u} p(f)df/(f_u - f_d) \tag{1}$$

In Formula (1): f_u is the upper limit of the Alpha waves frequency band; f_d is the lower limit of the Alpha waves frequency band; $p(f)$ is the power spectral density of the EEG signal.

$$R = G(\alpha)/G(\beta) \tag{2}$$

In Formula (2): R is the average power ratio of the Alpha waves and Beta waves.

According to the video and acceleration, the dangerous scenarios were extracted, and the brain wave data of 3 s after the occurrence of the dangerous scenarios was collected to analyze.

Since personal sign of each driver was various, EEG variation rates were adopted in order to reasonably measure the fluctuation of EEG caused by dangerous scenarios.

$$VAR = \frac{R}{R_r} \tag{3}$$

In Formula (3), R is the value of the EEG index under various dangerous scenarios, R_r is the reference value of the EEG indicator under the driver's steady state.

4 Analysis of EEG Variation Under Various Dangerous Scenarios

In order to analyze whether there is a significant difference in the impact of dangerous scenarios on EEG variation rates, the correlation analysis of EEG variation rates is performed by SPSS.

Correlation analysis using Pearson method

$$r_{xy} = \frac{\sum[(x_i - \overline{x})(y_i - \overline{y})]}{\sqrt{\sum\left[(x_i - \overline{x})^2(y_i - \overline{y})^2\right]}} \tag{4}$$

In formula (4), $\overline{x}$—the arithmetic mean of $x_i (i = 1, 2, \ldots, n)$, $\overline{y}$—the arithmetic mean of $y_i (i = 1, 2, \ldots, n)$.

According to the correlation analysis test result $P = 0.00000013 < 0.05$, the significance level is $\alpha = 0.05$. It can be known that the original hypothesis is rejected, and the dangerous scenarios types have a significant impact to the driver's EEG variation rates.

Figure 2 shows that the EEG variation rates line diagram of 14 drivers. The severity of the impact on the EEG variation rates is from strong to weak: Pedestrian (bicycle) conflict without prior vehicle maneuver, Front car decelerating suddenly, Front car off lane, Pedestrian (bicycle) conflict with prior vehicle maneuver, Vehicle changing lane. Further, the EEG variation rates mean values under the five various dangerous scenarios were 4.501, 4.105, 3.958, 3.660, and 3.349.

Comparing the above-mentioned EEG variation rates in different dangerous scenarios, the analysis found that when there is a potential risk including pedestrian or a non-motor vehicle such as a bicycle, the driver's EEG variation rates changes greatly due to directly exposed to traffic by pedestrians or bicycle riders. It shows that the driver is more nervous when facing vulnerable road users, and has a greater driving workload when the danger occurs. When the vehicle turns to meet the pedestrian, EEG variation rates is smaller than go straight the pedestrian (bicycle). The driver who turns at the intersection has a lower speed and is mentally prepared to observe the dynamic movement of pedestrians and other traffic participants. For three dangerous scenarios involving motor vehicles, there is a potential risk of collision before the sudden deceleration of the preceding vehicle because the driver does not have psychological expectations for the occurrence of danger. Therefore, when the danger occurs, the driver's EEG variation rates changes significantly, which is the most

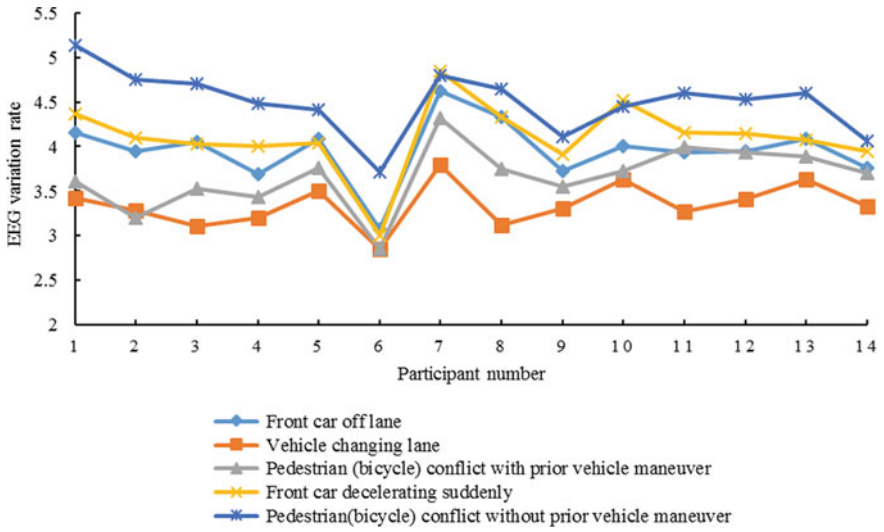


Fig. 2 The EEG variation rates of various dangerous scenarios

severe compared with others. Under this scene, the driver is extremely nervous with high driving workload. When other vehicles invade the lane, the intruding vehicle driver will drive reasonably to minimize the impact on the disturbed vehicle, and the driver is in a state of tension. The EEG variation rates of the vehicle's lane change to other vehicles is smaller because the driver is ready to change lanes in advance, causing low driving workload.

5 Analysis of Factors Affecting Driver's EEG Variation

Based on past research results and actual driving experience, the driver's proficiency in vehicle handling, age and other factors will have a certain impact on the driver's driving workload under dangerous scenarios. The following part is the influence of the driving experience and age of driver on EEG variation.

5.1 The Driving Experience

Drivers are categorized into inexperienced drivers and experienced drivers based on their driving years. The driving time of inexperienced drivers less than 8 years and the driving time of experienced drivers are more than 8 years. A comparative analysis of the EEG variation rates plots for each type driver under various dangerous scenarios was performed as well as Pearson correlation. $P = 0.000083 < 0.05$, $\alpha = 0.05$

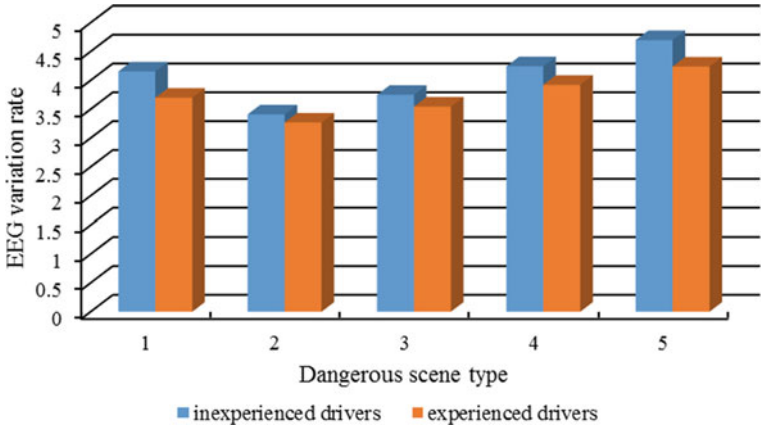


Fig. 3 The EEG variation rates of drivers with different driving experience Tips: (1) Front car off lane. (2) Vehicle changing lane. (3) Pedestrian (bicycle) conflict with prior vehicle maneuver. (4) Front car decelerating suddenly. (5) Pedestrian (bicycle) conflict without prior vehicle maneuver

indicating that driving experience has a significant impact on the EEG variation rates.

Figure 3 illustrates that the driver's EEG variation rates vary significantly with driving experience under various dangerous scenarios. The EEG variation rates of the experienced driver's EEG indicator is smaller under five dangerous scenarios while inexperienced driver is larger, indicating experienced driver's strong ability to deal with dangerous scenarios. The reason why they can process dangerous scenarios well is that the experienced driver has a relatively long driving time and has opportunity to learn from experience. Therefore, with the occurrence of dangerous scenarios, experienced drivers perform faster and more accurately with smaller driving load, while the inexperienced driver is likely to be confused and nervous with a larger driving workload due to insufficient driving experience.

5.2 The Age of Driver

According to age from young to old, drivers are divided into three groups: youth (20–34 years old), middle-aged (35–49 years old) and old (50–60 years old), and the test data of each group are grouped. Table 4 shows the EEG variation rates mean of drivers of different ages under various dangerous scenarios. Pearson correlation analysis was performed. $P = 0.0039 < 0.05$, $\alpha = 0.05$ indicating that age has a significant effect on the EEG variation rates of drivers. It can be seen that the EEG variation rates of the middle-aged group is the smallest, and the EEG variation rates of the young group is largest, and the old group is higher than the middle-aged group. The analysis shows that the driving workload of young driver is the greatest due to insufficient driving

Table 4 The EEG variation rates mean for different age groups

Dangerous scenarios	Youth	Middle-aged	Old
Front car off lane	4.129	3.796	3.789
Vehicle changing lane	3.449	3.206	3.278
Pedestrian (bicycle) conflict with prior vehicle maneuver	3.748	3.445	3.670
Front car decelerating suddenly	4.229	3.949	4.003
Pedestrian (bicycle) conflict without prior vehicle maneuver	4.706	4.002	4.315

experience and unstable and immature status, which means they are vulnerable to dangerous scenarios. Although older drivers' operational responsiveness tends to weaken with age, older drivers have more driving experience. Hence, their EEG variation rates are slightly higher than the middle-aged group as well as slightly higher driving workload. The middle-aged drivers have a wealth of driving experience so they can smoothly respond to various dangerous scenarios with a smallest driving workload.

6 Conclusion

(1) Various dangerous scenarios have different effects on the driving workload. When there exists dangerous objects easily caused serious traffic accidents such as pedestrians or bicycles and other non-motor vehicles and other targets, a greater impact on the EEG variation rates of driver is verified with appearances of higher driving workload. When there exists dangerous objects caused by external factors such as turning and encountering pedestrians (bicycles) or a relatively sound protection measures taken like vehicle interactions, the impact on the EEG variation rates of driver is much smaller with appearances of lower driving workload.
(2) Driving experience and driver age can affect driving workload. The longer the driving years and the richer the driving experience means the lower the EEG variation rates, indicating lower degree of tension when encountering dangerous scenarios. That is why inexperienced drivers with shorter driving years have higher EEG variation rates and driving workload. By contrast, middle-aged drivers have the lowest EEG variation rates with smallest driving workload while young drivers have the largest driving workload.

Acknowledgements This research has been supported under the National Key Research and Development Project (2017YFC0803901).

References

1. Agbayani H, Outrage A (2016) Assessment of mental workload by EEG + FNIRS. Conf Proc IEEE Eng Med Biol Soc, 3773–3776
2. Andrei GL, Takashi S (2011) Brain waves measurement based evaluation of mental workload related to visual information while driving. SAE Int J Passenger Cars—Mech Syst 4(1):578–585
3. Lim WL, Sourina O, Liu Y et al (2015) EEG-based mental workload recognition related to multitasking. In: 2015 10th international conference on information, communications and signal processing (ICICS). IEEE, New York
4. Kim IH, Kim JW, Haufe S et al (2015) Detection of braking intention in diverse situations during simulated driving based on EEG feature combination. J Neural Eng 12(1):016001
5. Sega S, Iwasaki H, Hiraishi H et al (2011) Verification of driving workload using vehicle signal data for distraction-minimized systems on ITS. In: 18th ITS World Congress Trans Core ITS America ERTICO-ITS Europe ITS Asia-Pacific
6. Kim JY, Jeong CH, Jung MJ et al (2013) Highly reliable driving workload analysis using driver electroencephalogram (EEG) activities during driving. Int J Automotive Technol 14(6):965–970
7. Kim HS, Hwang Y, Yoon D et al (2014) Driver workload characteristics analysis using EEG data from an urban road. IEEE Trans Intell Transp Syst 15(4):1844–1849
8. Hu J, Wang R (2017) Classification of driving workload affected by highway alignment conditions based on classification and regression tree algorithm. Traffic Inj Prev 19(1)
9. Meng W (2013) Study on driving workload characteristics of highway ramp area
10. Yuan W, Guo Y, Fu R et al (2014) Influence of urban road section types on drivers' workload. J Chang'an Univ (Natural Science Edition) 34(05):95–100
11. Wang C, Lu Y, Fu R et al (2016) Study on the physiological load characteristics of drivers when pedestrians suddenly invade the road. Sci Technol Eng 16(21):106–111

Classification of Beijing Metro Stations Based on Multi-source Data and Gaussian Mixture Model

Feng Wan, Jianrui Miao and Shuling Wang

Abstract At present, the metro plays an important role in people's daily travel. In order to clarify the function of the metro stations and to improve the service level of the metro, the reasonable classification of metro stations is particularly necessary. In this paper, multi-source data including Internet data and ridership data is obtained, and the data is analyzed to obtain 12 clustering initial variables. After that, 3 common factors are extracted from the 12 initial variables by factor analysis. According to the extracted common factors, 249 metro stations in Beijing are divided into 4 clusters by Gaussian mixture model, and the probability values that a station belongs to each cluster are obtained.

Keywords Classification · Multi-source data · Gaussian mixture model · Metro station · Beijing

1 Introduction

The metro station is a key node in the urban rail transit network and a gathering place for various social and economic activities in the city. There are differences in traffic functions and land-use functions of different types of stations [1]. At present, the classification of metro stations has not yet formed a unified standard. For example, Yu Lijie divided the 17 stations of the Xi'an Metro Line 2 into 5 categories based on the spectral clustering method [2]. Li Xiangnan selected 11 factors as the initial variables of cluster analysis and finally divided the 16 stations of the Chengdu Metro

F. Wan · J. Miao (✉)
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
e-mail: jrmiao@bjtu.edu.cn

F. Wan
e-mail: 17120878@bjtu.edu.cn

S. Wang
Beijing Transport Institute, Beijing, China
e-mail: wangshuling@bjtrc.org.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_88

Line 1 into 5 categories by K-means [3]. Yin Qin extracted 10 passenger flow features by using subway swipe data and divided 195 Beijing metro stations into 8 types [4]. According to the existing researching state, most scholars classify the stations through field survey data or land-use around the stations, and the methods are mostly hard clustering model. However, these data are difficult to obtain and the classification is too absolute.

In the context of rapid development of big data, the mining and analysis of multi-source data provides a new direction for research on classification of metro stations.

2 Data and Variable

2.1 Data Collection

Multi-source data consists of Internet data and ridership data. In this paper, Internet data includes the data from the real estate website and the Amap. Ridership data comes from the AFC (Automatic Fare Collection) data of Beijing Rail Transit from August 19 to 25, 2018.

The Internet contains a lot of valuable public data, and the web crawler technology is one of the important means to obtain the Internet data. A Web crawler is also known as a Web spider or Web robot [5], which is a program that initiates a request to a website to analyze and extract useful data after obtaining resources. The main steps for Web crawler to get the data are shown in Fig. 1.

The Amap JS API is a map application programming interface developed by the JavaScript, which provides many open services such as POI (Point of Interest) search

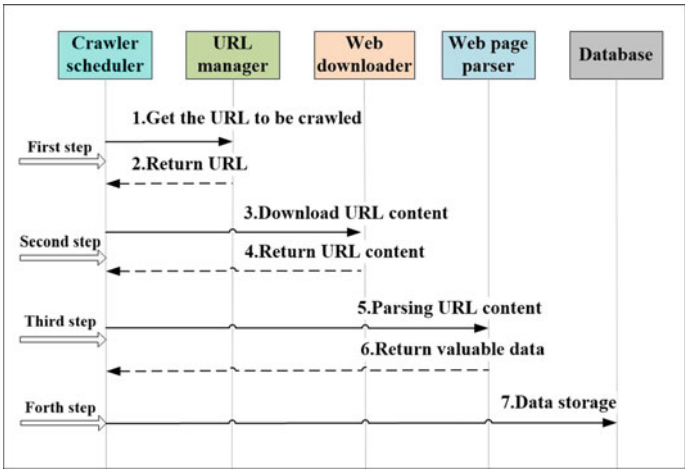


Fig. 1 Flow chart of the crawler

and route planning. The AFC data includes information such as the card number, the boarding (alighting) station and the inbound (outbound) time, which can describe the attributes of passenger's travel.

In order to obtain the multi-source data, the following major steps are performed. First, the crawler program is written in Python 3.6 to obtain data related to the residential area and the office building on the real estate website. Then, program is written to call the API interface of Amap to obtain the data of bus stations and various POI. After that, the AFC data is processed by using the Python programming to obtain the ridership data.

The Python data analysis packages and ArcGIS10.0 are used to analyze the Internet data and the ridership data, in order to form a dataset with the attributes of metro stations. Since the Internet data obtained is mainly data of residential and office types, the stations in transportation hubs and tourist attractions are not within the scope of this study. In summary, the dataset of 249 Beijing metro stations is finally obtained.

2.2 Variable Selection

The land-use around metro stations is the source of ridership, and the characteristics of stations determine the attraction intensity to the ridership. Therefore, this paper establishes the first-level indicator system from the two dimensions of the station characteristics and the land-use characteristics. The former includes the scale of passenger flow, number of feeder bus and the accessibility of metro station, etc. The latter includes the price and rent of residential buildings, area of office buildings and number of various POI, etc. After analysis of the first-level indicators, 12 secondary indicators are selected as the initial variables, which mainly include:

(1) Station characteristics: average inbound volume of morning-peak hour at weekday (S1), average outbound volume of morning-peak hour at weekday (S2), average inbound volume of evening-peak hour at weekday (S3), average outbound volume of evening-peak hour at weekday (S4), the number of bus stops within walking distance (S5), the number of other stations in the urban rail transit network that a station can reach within 20 min (S6). The morning-peak hour and the evening-peak hour mentioned above are 7:00–9:00 am and 17:00–19:00 pm respectively.
(2) Land-use characteristics: average rent of residential buildings (L1), average price of residential buildings (L2), total floor area of office buildings (L3), the number of business hotels, residential areas (L4), the number of schools, hospitals (L5), ratio of road length to the area (L6). These data are collected within the walking distance of metro stations. In this paper, the walking distance is 800 m.

The sample data of the 12 initial variables are shown in Table 1.

Table 1 Summary of variables

Variables	Mean	Standard deviation	Min	Max
S1	3476.93	2701.83	173	15,798
S2	2785.84	2873.79	60	16,757
S3	3640.86	4015.24	62	23,889
S4	2626.72	1954.78	144	11,736
S5	26.57	14.78	0	84
S6	14.99	8.49	1	38
L1	98.01	30.24	35.11	176.78
L2	70,558.50	23,354.12	32,465.16	144,279.38
L3	721,721.05	1,238,243.35	0.00	10,435,683. 7
L4	61.09	37.67	0	236
L5	269.38	197.55	0	1101
L6	7.02	2.70	0.39	13.38

3 Gaussian Mixture Model

3.1 Model Description

The GMM (Gaussian mixture model) is a linear combination of a finite number of Gaussian distributions [6]. Assuming a K Gaussian distribution, the GMM formed by mixing the K Gaussian distributions with a certain probability is as follows:

$$P(x) = \sum_{k=1}^{K} p(k) \cdot p(x|k) = \sum_{k=1}^{K} \pi_k \cdot N(x|\mu_k, \Sigma_k) \tag{1}$$

where, π_k, μ_k, Σ_k are probability, mean and variance of the kth Gaussian distribution respectively, $N(x|\mu_k, \Sigma_k)$ is the probability density function of the kth Gaussian distribution, the formula is:

$$N(x|\mu_k, \Sigma_k) = \frac{1}{\sqrt{2\pi|\Sigma_k|}} \exp\left(-\frac{(x-\mu_k)^2}{2\Sigma_k}\right) \tag{2}$$

The probability density function of the GMM is:

$$f(x) = \sum_{k=1}^{K} \pi_k f_k(x; \theta_k) \tag{3}$$

where, f_k is the distribution density of the kth Gaussian distribution, π_k and θ_k are parameters and there are $\pi_k \in [0, 1]$ and $\sum_{k=1}^{K} \pi_k = 1$.

When the GMM is used for clustering, the K Gaussian distributions represent that the final clustering number is K. Both this model and the K-Means model belong to the unsupervised learning model in machine learning algorithms. The difference is that the GMM belongs to the probability model. The final clustering result of each data is a probability value uniquely generated by K Gaussian distributions, rather than being absolutely belonging to a certain class, that is, each Gaussian distribution can generate this data, but the probability of occurrence is different.

3.2 Steps of Clustering by GMM

The main steps of clustering based on GMM are as follows:

(1) For data x_i, the probability generated by the kth Gaussian distribution is as follows:

$$\psi(i,k) = \frac{\pi_k \cdot N(x_i|\mu_k, \Sigma_k)}{\sum_{n=1}^{K} \pi_n \cdot N(x_i|\mu_n, \Sigma_j)} \tag{4}$$

(2) The VBEM (Variational Bayesian Expectation Maximization) algorithm is used to estimate the model parameters. In order to effectively cluster the data $x_i (i = 1, 2, \ldots, n)$, the model needs to be parameterized to maximize the probability of generating this data. The maximum likelihood function L is:

$$L(\pi_k, \theta_k) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \sum_{k=1}^{K} \pi_k f_k(x_i; \theta_k) \tag{5}$$

The parameters are estimated by the VBEM algorithm until the model converges to obtain stable results. The advantage of this algorithm is that it can automatically select the appropriate K value, helping us to quickly determine the final number of clusters.

4 Clustering Results

The basic data of clustering in this paper is the dataset of 249 Beijing metro stations which includes 12 initial variables. The method of clustering is GMM.

4.1 Factor Analysis

Since there may be strong correlation between the 12 variables obtained above, it is necessary to perform correlation analysis on them. When the correlation coefficient is closer to 1 (−1), it means that the positive (negative) correlation between the two is stronger; when the correlation coefficient is 0, it means that the two are irrelevant. Figure 2 shows the correlation coefficient of the initial variables. It can be seen from the figure that some variables have a high correlation.

Due to the strong correlation between some variables, the data needs to be reduced in dimension. Factor analysis is used to extract hidden common factors that can explain the original variables [7], which can effectively eliminate the influence of data collinearity and improve the accuracy of classification. Firstly, KMO (Kaiser-Meyer-Olkin) and Bartlett's test are performed to examine the structural validity of the data. The results are shown in Table 2, where the KMO's coefficient (0.803) is greater than 0.0 and the significance (0.000) is less than 0.001, indicating that the dataset is suitable for factor analysis.

The extraction results of the factor analysis are shown in Fig. 3. The extraction ratio of 12 initial variables are all above 57%, indicating that the common factors can effectively extract the information of the initial variables.

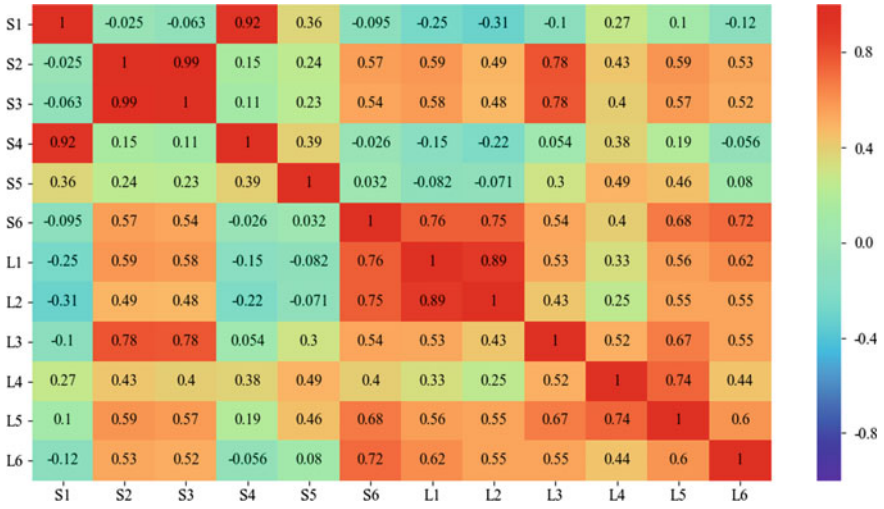


Fig. 2 Correlation coefficient of initial variables

Table 2 KMO and Bartlett's test

KMO measure of sampling adequacy	Bartlett's test of sphericity		
	Approx. Chi-Square	df	Sig.
0.803	3390.619	66	0.000

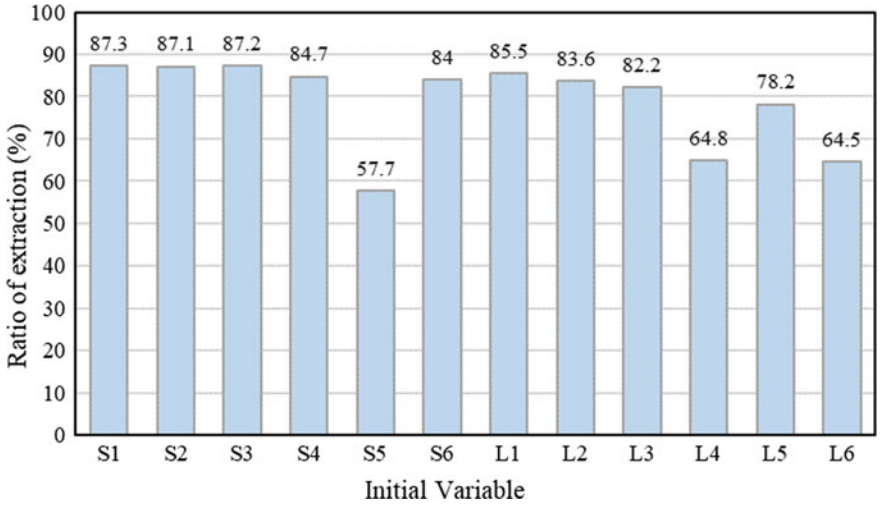


Fig. 3 Ratio of extraction

From the scree plot, we can see that when the number of common factors are equal to 3, the graph tends to be horizontal, so we can extract 3 main common factors from these initial variables (Fig. 4).

According to Table 3, it is found that factor one is the embodiment of land-use characteristics, which has high load on variables S6, L1, L2, L5, L6. While factor

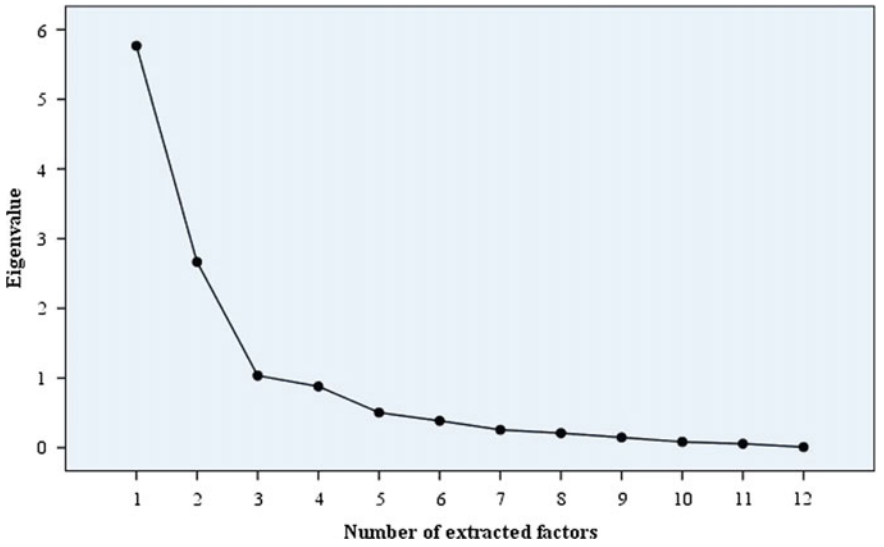


Fig. 4 Scree plot

Table 3 KMO and Bartlett's test

Factor	High load variable	Factor name
Factor one	S6, L1, L2, L5, L6	Land-use factor
Factor two	S2, S3, S5, L5	Outbound volume of morning-peak hour factor
Factor three	S1, S4, L3, L4	Inbound volume of morning-peak hour factor

two and factor three are the embodiment of station characteristics, since the two have high load on variables related to the ridership value.

From Table 4, the 3 common factors with eigenvalues greater than 1 can explain 78.899% of the information of the original sample, indicating that the 3 common factors can basically replace the 12 initial variables. In summary, the results of the factor analysis are reasonable.

4.2 Results Analysis

Based on the results of factor analysis, take the 3 common factors as input and the clustering results as output. Finally, 249 metro stations in Beijing were divided into 4 clusters by GMM. Some station names are abbreviated as follows (Table 5).

The top 5 stations with the highest probability of each cluster are shown in Table 6. The probability belonging to each type and the geographic location of stations are visualized by ArcGIS Pro, as shown in Fig. 5. The clustering results are analyzed and the following conclusions are obtained:

(1) The stations with high probability of type one are mainly located in typical office areas such as GuoMao, Financial Street and ZhongGuanCun, including GM Station, FXM Station and ZGC Station, etc. There are large-scale office land-use around these stations, and the nature of the land-use is relatively single, so the stations of type one are defined as the large-scale employment stations.
(2) The stations with high probability of type two are mainly located between the South Second Ring Road and the North Fifth Ring Road, including RJDJ Station, WJS Station and BXQ Station, etc. The surrounding areas of these stations are mainly office land-use, but the scale and the purity of land-use are not as good as the stations of type one, so the stations of type two are defined as the general employment stations.
(3) The stations with high probability of type three are mainly located in typical residential areas such as HuiLongGuan, TianTongYuan and SongJianZhuang, including HLG Station, TTY Station and SJZ Station, etc. There are large-scale residential land-use around these stations, and the nature of the land-use is relatively single, so the stations of type three are defined as the large-scale residential stations.
(4) The stations with high probability of type four are mainly located in the suburban metro lines outside the South.

Table 4 Total variance explained

Component	Initial eigenvalues			Extract square and load			Rotation square and load		
	Total	Variance (%)	Cumulative (%)	Total	Variance (%)	Cumulative (%)	Total	Variance (%)	Cumulative (%)
1	5.768	48.067	48.067	5.768	48.067	48.067	3.690	30.751	30.751
2	2.665	22.212	70.279	2.665	22.212	70.279	3.221	26.844	57.594
3	1.034	8.620	78.899	1.034	8.620	78.899	2.557	21.305	78.899
4	0.879	7.325	86.224						
5	0.505	4.205	90.429						
6	0.386	3.214	93.643						
…									
11	0.057	0.475	99.929						
12	0.009	0.071	100.000						

Table 5 Abbreviations of partial station names

Station name	Abbr.	Station name	Abbr.
ZhongGuanCun	ZGC	HuoYing	HY
YongAnLi	YAL	HuiLongGuan	HLG
XiErQi	XEQ	ZhangGuoZhuang	ZGZ
GuoMao	GM	FengBo	FB
FuXingMen	FXM	DaoTian	DT
RongJingDongJie	RJDJ	LinHeLi	LHL
WangJing South	WJS	NanShao	NS
TsinghuaDongLuXiKou	TDLXK	WanYuanJie	WYJ
CheDaoGou	CDG	BeiShaTan	BST
GuangQuMenNei	GQMN	ZaoYing	ZY
TianTongYuan North	TTYN	JinSong	JS
TianTongYuan	TTY	BaiShiQiao South	BSQS
LiShuiQiao	LSQ	XiZhiMen	XZM

Table 6 Clustering results of partial stations

Clustering	Station	Probability value (%)			
		Type one	Type two	Type three	Type four
Type one	ZGC	100.0	0.0	0.0	0.0
	YAL	100.0	0.0	0.0	0.0
	XEQ	100.0	0.0	0.0	0.0
	GM	100.0	0.0	0.0	0.0
	FXM	100.0	0.0	0.0	0.0
Type two	RJDJ	0.5	97.3	0.2	2.1
	WJS	2.4	97.2	0.3	0.1
	TDLXK	2.1	97.1	0.5	0.3
	CDG	2.8	96.8	0.2	0.2
	GQMN	2.7	96.7	0.0	0.6
Type three	TTYN	0.0	0.0	100.0	0.0
	TTY	0.0	0.0	100.0	0.0
	LSQ	0.0	0.0	100.0	0.0
	HY	0.0	0.0	100.0	0.0
	HLG	0.0	0.0	99.9	0.1
Type four	ZGZ	0.0	0.0	0.0	100.0
	FB	0.0	0.0	0.0	100.0
	DT	0.0	0.0	0.0	100.0
	LHL	0.0	0.0	0.1	99.9
	NS	0.0	0.1	0.0	99.9

(a) Type one (b) Type two (c) Type three (d) Type four

Fig. 5 Probability and location of the stations belonging to each type

Third Ring Road and the North Fifth Ring Road, including ZGZ Station, FB Stations and DT Stations, etc. The surrounding areas of these stations are mainly residential land-use, but the scale and the purity of land-use are not as good as the stations of type three, so the stations of type four are defined as the general residential stations.

Some stations have complex land-use around them, and these stations do not strictly belong to one of the types mentioned above. Clustering based on GMM can calculate the probability value that a metro station belongs to a certain type, which represents the situation of land-use around the station. Therefore, compared with hard clustering such as K-means, GMM has advantages in depicting the land-use mix and land-use scale around the metro stations. The following table lists some stations with complicated land-use.

From Table 7, there are both residential and office land around WYJ Station, BST Station and ZY Station, and the land-use scale of the two types are small. The proportion of residential land (68.8%) around WYJ Station is larger than that of office land (31.0%), while that around ZY Station is just the opposite. The ratio of the two types of land-use around BST Station is not much different. The scale of residential land around JS Station is larger than that of general residential stations, but not as far as large-scale residential stations, such as HLG Station and TTY station. The scale of office land around BSQS Station is larger than that of the general employment

Table 7 Probability of partial comprehensive stations

Station	Probability value (%)			
	Large-scale employment station	General employment station	Large-scale residential station	General residential station
WYJ	0.1	31.0	0.1	68.8
BST	0.6	53.0	0.1	46.2
ZY	1.2	77.6	0.0	21.2
JS	1.0	1.0	24.9	73.1
BSQS	20.4	79.5	0.2	0.0

stations, but it is inferior to the large-scale employment stations such as GM Station and ZGC Station.

5 Conclusion

In the context of the development of big data, this paper obtains the multi-source data by various methods. Then, the dataset of 249 stations in Beijing Metro is obtained and 12 initial variables are selected. Through data mining and analysis, 3 common factors are extracted. Finally, based on the common factors and GMM, 249 stations are divided into 4 clusters: large-scale employment station, general employment station, large-scale residential station and general residential station. By analyzing the land-use characteristics around stations of each type, it is found that GMM can not only classify the stations reasonably, but also well describe the land mix and the land-use scale by the probability that a station belongs to each type.

This paper provides new data source and feasible method support for related research. However, there are still some shortcomings in data acquisition. The data of transportation hubs and tourist attractions are not taken into consideration. In the follow-up study, the database needs to be further enriched to achieve a more comprehensive research.

Acknowledgements The study is supported by the project that completed by State Key Laboratory of Rail Traffic Control & Safety and Beijing Transport Institute. The generous assistance of professor Miao is greatly appreciated.

References

1. Duan D, Zhang F (2013) Research on urban orbit site classification from the perspective of land use optimization—Taking Xi'an Metro Line 2 as an example. Urban Plann 37(09):39–45
2. Yu L, Li Y, Chen K (2014) Classification method of urban orbital sites based on spectral clustering. J Traffic Inf Safety (1)
3. Li X (2015) Study on clustering methods of urban rail transit site classification. Railway Standards 59(04):19–23
4. Yin Q, Meng B, Zhang L (2016) Identification of Beijing metro station types based on passenger flow characteristics. Prog Geogr 35(01):126–134
5. Kumar M, Bindal A, Gautam R et al (2018) Keyword query based focused web crawler. Proc Comput Sci 125:584–590
6. Chen Y (2016) Variable selection and application based on Gaussian mixture model clustering
7. Kim KW, Lee DW, Chun YH (2010) A comparative study on the service coverages of subways and buses. KSCE J Civil Eng 14(6):915–922

The Progressivity of a Per-kilometer Congestion Tax in Beijing

Tian Yu

Abstract In recent years, the traffic congestion problem in Beijing during peak hours has attracted more and more attention. And a series of charging policies or researches were designed to relieve the congestion. In this article, we focus on the per-kilometer congestion tax and carry on empirical analysis to measure its progressivity. Based on the survey data,we calculate the mileage of each sample, add them up, and divide these samples into different groups according to income level and spatial district. Then we introduce the Suit index to analyze income and spatial distributional effects of the per-kilometer congestion tax among different groups. It is found that a per-kilometer congestion tax is regressive in terms of the income distribution. That is, compared to higher income group, the lower would pay more for the tax. And in the case of spatial distribution, the tax is progressive. That is, compared to urban counterparts, the rural residents would bear more tax burden. Finally, the econometric multiple regression method was applied to explain the characteristics of resident commuting mileages in Beijing. The result suggests that the highly educated residents tend to travel far for work.

Keywords Congestion fee · Suits index · Tax progressivity · Spatial mismatch

1 Introduction

Traffic congestion is a recognized problem in urban development. Congestion also exists in large and medium-sized cities in China, especially in the key roads during peak hours. It has been shown that 61% Chinese cities' traffic are in a slow speed and 13% are in congested state during peak commuting hours in 2018. At present, the traffic speed in CBD during peak hours is only 15–20 km [1].

Traffic congestion remains a particularly serious problem in Beijing. According to the Amap traffic data, the delay index of Beijing road network in peak hours is 2.032, ranking first among 50 domestic cities. And the congested accounts for 11.08% of

T. Yu (✉)
School of Economics and Management, Beijing Jiaotong University, Beijing, China
e-mail: 17120531@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_89

all network in peak hours. That is, 11.08 per 100 km are in the congestion even serious congestion condition, which is also the most severe compared to other cities. In Beijing, the average travel delay caused by congestion is 44.97 min every day. In one year (232 working days), the annual delay time is 174 h, which is equivalent to 22 working days, or 8400 yuan economic loss [2].

Meanwhile, the number of motor vehicles in Beijing had reached 59.09 million in 2017, an increase of 3.4% over the previous year. And the number of private vehicles had reached 47.56 million, an increase of 3.3%. And the applicants for license plate continue to increase [3]. Although the government has implemented relevant measures to control the total vehicles number and the growth rate has slowed down, the number of vehicles itself in Beijing is large enough. And the residents' travel demand, especially the commuting demand during peak hours, has increased, not decreased. Therefore, Beijing's urban road network is under tremendous pressure, and congestion has become an urgent problem to be solved for the sustainable urban traffic development.

Internationally, congestion pricing policy has been proved and applied as an effective measure to alleviate road congestion. Singapore launched the first congestion toll system in 1975, and began to set up electronic toll system in 1998, which charges vehicles by times. After the congestion charge was implemented, the traffic flow decreased by 45%, and the average traffic speed increased from 18 to 35 km/h during peak hours [4, 5]. Another example is the London. Vehicles travelling through 22 km^2 of the city center have to charge at a price of 5 lb per day during a fixed period of time. Vehicles paying the fee can enter the toll area several times in one day [6]. After that, the average speed increased from 5 to 20 km/h, and the congestion level decreased by 30% [4].

In China, congestion pricing issue has been involved in many research fields or policy experimentation. However, congestion pricing policy has not been formally implemented. In this context, referring to international practical examples and theoretical frontiers, this paper studies the equity concerns and distribution effects of the per-kilometer congestion tax in Beijing, combining with the current commuting situation of urban residents, which can provide insight for the formulation and implementation of congestion pricing policy in China.

2 Literature Review

Pigou has long proposed the concept of road tolls in welfare economics to prove the rationality of congestion charges. He believes that the social negative externality caused by car users can be solved by the Pigou tax levied by the government, which is an effective policy to curb congestion. Afterwards, in 1969, Vickrey put forward the famous bottleneck model and concluded that congestion pricing can eliminate the queuing in the bottleneck [7]. In China, with the growth of urban traffic congestion problem, economists have tended to focus on the congestion pricing research. There is relevant study combining the change of congestion level and welfare of different

people groups. It is pointed that the reasonable congestion fee can curb congestion and improve the overall social utility [8]. Considering environmental cost caused by travel, Zhao Hongjun et al. adopted an economic model, and found that congestion pricing is an effective way to bridge the gap between private driving costs and social cost, which would eventually lead to people travel mode change, and the environmental cost internalization [9]. Using an impedance function and a traffic assignment model, Zhao Ying et al. formulated a virtual toll scheme for a specific road section in Nanjing, evaluated indices such as the traffic volume and travel mode split rate, and reached the conclusion that congestion pricing can have valid improvement effects [10].

Lots of countries have explored themselves in the pricing patterns and charging rate, because of the differences in the urbanization level, spatial structure, road planning, technology and so on.

In the theoretical field, XIAO studied a tactical queuing problem under flat toll, and commuters are faced with discontinuous travel costs. The optimal toll can eliminate queuing delay to some extent. The article also points out the flat toll has drawbacks itself, which can be reflected in the welfare loss of all commuters and the problem that the revenue is not enough to pay for the road construction [11]. In recent years, on the basis of constant and linear residents' utility, Chinese scholars extended Vickrey's bottleneck model to an activity-based bottleneck model to address the departure time choice and obtained the optimal solution of time-varying toll and step toll [12]. There are also some studies based on the commuter time value and vehicle emissions, proposed congestion pricing strategies, and evaluated the toll area, tax rate, technical feasibility and other issues [13, 14]. The virtual toll scheme in Nanjing showed that the improvement effect of congestion toll is not in directly proportional to the charging rate and the rate corresponding to the best improvement effect exists [10].

In the practical field, congestion pricing measures have been implemented and achieved good results. As mentioned above, the electronic toll system in Singapore can automatically scan and charge passing vehicles; in London, one vehicle with a pass can enter or leave the toll area without restriction. In addition, Oslo, Stockholm, Melbourne, Stuttgart, Toronto and other cities have also adopted different congestion pricing schemes.

Up to now, no country has adopted mileage-based congestion tax, but in the Oregon, United States, has experienced success in pioneering mileage-based fees. To compensate for the traffic budget, in 2015, Oregon government assessed a charge of 1.5 cents per mile instead of the traditional fuel tax for vehicle. Essentially, it belongs to the reform of road usage charging system, not congestion charges, but through the practice, the U.S. government has found the advantages of mileage-based fee. That is, for consumers, mileage-based fee is more direct and the revenue would not decrease with the increase of vehicle efficiency. And the extra fee in specific areas or peak periods can achieve public transport increase and congestion alleviation.

Many scholars presented their own views on the advantages of mileage-based fee. Ke and Gkritza [15], using the data from the Oregon Department of Transportation,

through empirical research, found the per-mile congestion tax could save the administratively redistribution policy. And it avoids geographic problem of choosing one specific area to toll, but involves all traffic flow during peak hours regardless of drivers' routes. Moreover, to maximize the usage of price signal to guide travel behavior, different pricing rate should be applied between high-frequency, long-distance and low-frequency, short-distance vehicles, which exactly can be achieved by per-kilometer tax [16]. Considering the advantages and prospects of mileage-based pricing, this paper focuses on the per-kilometer congestion tax.

The equity issue of tax has always attracted wide attention. There are also many researches or studies about the effects of different tax policies from various angles. Policymakers are also worried about the possible regressive nature of congestion tax [17], which means the tax burden borne by low-income group accounts for a larger proportion of their income, compared to high-income group. Based on the data from European cities, Eliasson [18] analyzed the congestion tax effects from the perspectives of consumers and residents, and proved it is regressive. Krol [17] reviewed the traffic tax measures in U.S., then explored it in economic view. Just like Eliasson, he found both congestion tax and fuel tax are regressive. Furthermore, Ke and Gkritza [15] extended the income progressivity of congestion tax to spatial distribution. Among the area of different degree of rurality, tax is found to be regressive. This paper examines the progressive nature of a congestion tax in the specific Chinese city background. We study the equity concern of per-kilometer congestion tax in our country.

3 Data and Methods

3.1 Data

The data used in the study is from empirical survey. The original data is non-aggregated data, mainly coming from the results of the questionnaire in 2018 and 2019. To collect enough data, we made investigation in some parking lots, metro and bus stations and shopping malls through face to face interviews and sending questionnaires.

The survey was conducted twice. Among the results, 604 valid questionnaires of 732 were backed in December 2018, and 995 valid questionnaires of 1231 were backed in February 2019. In total, we collected 1559 valid questionnaires. The questionnaire is designed for commuters, especially motorists. It involves questions about weekday travel, such as whether they travel during peak hours, which travel modes they use, as well as the household socio-economic information, such as monthly income, where their family are located. Among 1559 respondents, 933 choose "driving" or "taxi" as travel modes. Combined with the interviews, the total number of samples selected for this study is 1190.

In order to calculate the commuting mileage, we paid much attention on the commuting start point and end point. As the trend of home-work separation in Beijing has been noticed [1, 19, 20], and the separation degree is relatively high [20], considering current situation in Beijing, we selected several employment centers and residential centers as alternatives in the questionnaires for the respondents to choose. They also can fill in the address or mark it on the map.

3.2 Suits Index

Since Daniel B. developed the suit index, it has enjoyed wide application [21], and has played an important role in policy formulation. In recent years, Chinese scholar discussed the relative of several indices of progressivity measurement and clarified the concept of progressive tax [22, 23].

Suit index is used to measure tax progressivity. The process is similar to the Gini coefficient. But it compares a cumulative frequency distribution of tax burden with a similar distribution of household income. In the figure, points of "Lorenz curve" above the 45° line suggest lower-income group pay more than their proportion of total income, that is, the tax is regressive and the index sign is minus. And points below the line show low income group pay a less proportion of the tax than their proportion of total income, compared to high income group, suggesting a progressive tax and plus sign of the index. The 45° line represents the points where the proportion of the tax paid by each income group exactly equals their proportion of population. In that case, the Suit index is equals to 0.

We also applied the Suit index to analyze the spatial progressivity of tax. It also can be judged by the corresponding figure that rural households shoulder a disproportionate tax burden. The process is similar to the previous.

Mathematically, the Suit index is calculated as follows:

$$S = \sum_{i=1}^{n-1} \left[Y_i T(Y_{i+1}) - Y_{i+1} T(Y_i) \right]$$

where S is Suits index value, n is the number of income groups or location types, Y_i is the accumulated percent of income, including the i th income group, and T is their accumulated percent of tax burden.

4 Results and Discussion

Combined with the salary level in Beijing, income groups used in the study are 0–4000, 4000–8000, 8000–12,000, 12,000–16,000, and 16,000 yuan or more monthly.

To examine the distributional impacts by income group, the medians of each group are used. 18,000 is used as the median income for the 16,000 yuan or more income-group.

Table 1 shows the distribution of households by income group.

For each group, the cost of driving during peak hours can be calculated by multiplying the number of miles driven in peak hours by a congestion tax rate. The results from the survey show that nearly 31% of residents tend to pay less than 20 yuan for congestion fee, while less than 2% of respondents are willing to pay 40 yuan. Base on the weighted average calculation, the average congestion fee which can be accepted by most residents is 23.85 yuan. Considering the average commuting mileage of the samples in that survey is 25.4 km, the tax rate is arbitrarily set at 0.938 yuan/km. To simplify calculation, we adopted one yuan per kilometer. That is, the congestion fee during peak hours is equal to the mileage.

As the following figure shows, the tax is regressive and the Suit index value is − 0.208, which suggests the lower income group will shoulder more burden of the tax (Fig. 1).

The data from Beijing salary level report in 2018 suggested that residents whose salary is between 4500 and 8000 yuan account for 37.2% of the whole. In our survey, 32.1% respondents get salary of 4000 and 8000 yuan a month. It is shown that the number of residents belonging to the middle and low income groups is large. And the

Table 1 Number of samples by income group

Median income (yuan)	Total residents
2000	87
6000	279
10,000	530
14,000	183
18,000	111
Total	1190

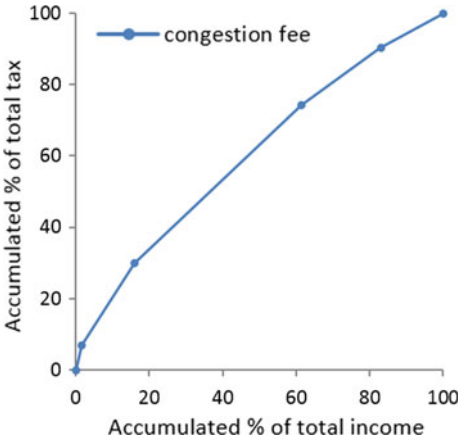


Fig. 1 Suit index of the congestion tax in terms of the income distribution

spatial mismatch should not be neglected in Beijing and the low-income are easily affected by urban structural changes [19]. They tend to commute longer under the pressure of vocational skills, housing price or children education, etc. Moreover, according to our survey, 85% households will own more than one car in the near future. And through cross analysis, we found 49% middle or low income households have had more than one car.

Therefore, low income group may bear the burden of the per-kilometer congestion tax disproportionately. The tax policy should be accompanied with revenue redistribution measures to balance the benefits of different income groups.

For the purpose of spatial distributional effect evaluation, we divide the city into 1–5 location types to describe the degree of rurality. The criterion is drawn from statistic yearbook of Beijing. The location type 1 contains Dongcheng and Xicheng district. The Chaoyang, Fengtai, Shijingshan and Haidian form location type 2. Fangshan, Tongzhou, Shunyi, Changping, Daxing district are coded as location type 3. And the other districts are the location type4. All respondents were asked which location type they live in (Fig. 2).

Table 2 shows the distribution of households by income group and location type.

Fig. 2 Location types in Beijing

Table 2 Number of samples by income group and location type

Median income (yuan)	Total residents	Location type 1	Location type 2	Location type 3	Location type 4
2000	87	5	31	35	16
6000	279	60	93	115	11
10,000	530	130	136	255	9
14,000	183	51	92	38	2
18,000	111	7	69	32	3

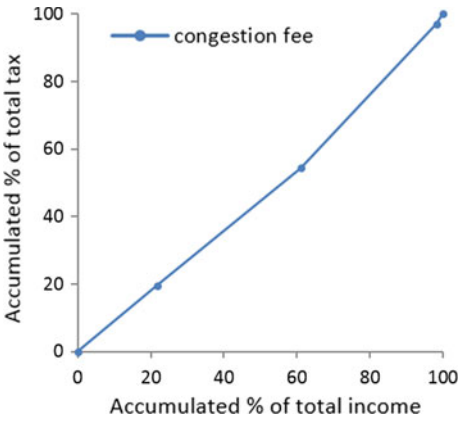


Fig. 3 Suit index of the congestion tax in terms of the spatial distribution

As the following figure shows, the tax is progressive and the Suit index value is −0.070. When considering spatial distribution effect, the rural residents will pay more for the tax than the urban (Fig. 3).

Compared to urban residents, the rural have to travel further by nature. And the suburbanization has contributed to the current job-housing location reconstruction in Beijing [24]. Huge residential districts such as Tiantongyuan, and Longze have emerged in suburb in the past few years. The results of the questionnaire show that 30% households are located in the "location type 3". The increase of population in suburb must be accompanied with the travel demand growth.

5 Regression Results

The average commuting mileage drawn from our data is 25.4 km, which is longer than any other city in our country. It also has been noted that the long time and far distance commuting is a common characteristic of residents in Beijing.

To better understand the factors influencing commuting mileage, regression analysis is used in this part to examine these socio-economic determinants. The dependent variable is the kilometer, the length of commuting. The independent variables are the commuters' gender, age, and education level, etc.

The regression results are shown in Table 3.

It is not hard to find that the commuters' education, income and number of household labors are more significant than other variables for all residents. As expected, lower income residents are more likely to be affected by socio-economic factors such as their gender, the number of household children and labor, to have a longer trip. This is due to there being more flexibility for higher income group. They can shift their job or housing location easily and avoid long-distance travel. We also can't deny the gender as a role in the commuting travel even the workplace. The male with higher income tend to driver more, while for the female, the increase in education is associated with the increase in mileage. Interestingly, the single people with older ages and Higher education often spend more time on commuting than others. For the married, the presence of a family car indicates a decreased likelihood of commuting too long. Through the positive estimated coefficients, we find high education level may encourage longer commuting on the premise that these highly educated people often seek suitable employments in broad space. They are cautious about career choice for influencing factors such as professional matching, personal interests and working environment, which inevitably causes longer commuting distance. While the low-educated people, mostly engaged in social services, choose jobs by proximity rather than interests or professions, leading to a shorter distance commuting.

6 Conclusions

This paper uses the Suit index to analyze the progressivity of a per-kilometer congestion tax in Beijing through empirical survey. We found the tax is regressive in terms of the income distribution. That is, compared to higher income group, the lower would pay more for the tax. And in the case of spatial distribution, the tax is progressive. That is, compared to urban counterparts, the residents in suburb would bear more tax burden. The final regression analysis shows Higher education may encourage longer commuting.

The drawback of the data used is that it is collected by survey, and through manual calculation. These samples could not cover all traveling behavior, so the representation is limited. Another limitation is the fixed tax rate. We got it through simple division, which is not accurate enough. The estimation of an optimal congestion tax is beyond the scope of the study.

Congestion charging is an effective way to alleviate urban congestion, which has been proved theoretically and practically. As the toll policy is related to the whole transportation system and the interests of all citizens, further research and experiment

Table 3 Regression results

Para meter	Regression model (1)		Regression model (2)		Regression model (3)		Regression model (4)		Regression model (5)		Regression model (6)		Regression model (7)	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Constant	3.089	0.085	3.241	0.189	3.363	0.344	2.731	0.167	3.092	0.111	2.827	0.154	3.350	0.275
Age	−0.001	0.002	−0.003	0.004	−0.007	0.017	0.005	0.004	−0.002	0.002	0.005*	0.004	−0.016	0.084
Car	−0.007	0.025	−0.049	0.054	0.097	0.175	0.042	0.061	0.007	0.029	0.058	0.054	−0.060**	0.034
Child	0.006	0.019	−0.054*	0.041	0.061	0.206	0.018	0.045	0.003	0.023	0.015	0.039	−0.026	0.028
Edu	0.036**	0.022	0.056	0.056	0.016	0.121	0.022	0.047	0.037***	0.016	0.087**	0.047	0.029	0.033
Gender	0.001	0.030	−0.186*	0.120	0.053	0.108	–	–	–	–	−0.001	0.068	0.008	0.052
Income	0.022	0.018	–	–	–	–	0.038**	0.022	0.014	0.020	−0.009	0.030	−0.011	0.024
Labor	0.027	0.022	0.060*	0.045	−0.049	0.081	0.038	0.042	0.026	0.026	–	–	–	–
Samples	704		195		208		321		383		177		546	

*The coefficients are the ones that are significant at the 90% confidence level
**The coefficients are the ones that are significant at the 95% confidence level
***The coefficients are the ones that are significant at the 99% confidence level

are also needed for a thoughtful and appropriate tax plan. And the characteristics of residents commuting behavior in Beijing should be paid more attention in the future work.

References

1. Zheng SQ, Xu YF, Zhang XN, Yu D (2015) Jobs-housing balance index and its spatial variation: a case study in Beijing. J Tsinghua Univ (Science and Technology) 55(04):475–483
2. Amap (2019) Chinese major urban traffic analysis report 2018, Beijing
3. Beijing Transport Institute (2018) Beijing transportation development annual report 2018, Beijing
4. Duan JY (2010) Discussion on congestion charge. China Water Transp 10(11):49–50
5. Shen M (2015) A study of international research implementation of road congestion charging & adaptation research in China. China Municipal Eng 5:8–11+93–94
6. Chen XH (2012) Successful strategies of British city traffic congestion charge and reference. Shanghai Urban Manage 21(01):55–57
7. Vickrey WS (1969) Congestion theory and transport investment. Am Econ Rev 34:414–431
8. Xu Z, Ou GL (2012) Theoretical basis and policy analysis on congestion charge. China Ind Econ 12:18–30
9. Zhao HJ, Feng SW (2015) How to effectively Govern Beijing's transport congestion? An economic appraisal based on a congestion charge model with consideration of environmental cost. Urban Dev Stud 22(12):101–110
10. Zhao Y, Lu J, Zhang WJ, Sun XL (2017) Improvement effect analysis of congestion pricing using Logit model. J Harbin Inst Technol 49(03):80–85
11. Xiao F, Shen W, Zhang HM (2012) The morning commute under flat toll and tactical waiting. Transp Res Part B: Methodol 46(10):1346–1359
12. Li ZC, Ding J (2017) An activity-based study on bottleneck model and congestion charging. J Manage Sci China 20(08):93–101
13. Zhu YZ (2015) Research on the traffic congestion charge under travel value of time. Beijing University of Technology, Beijing
14. Wu ZQ (2013) Road congestion pricing strategies based on vehicle emissions for Beijing. Beijing Jiaotong University, Beijing
15. Ke Y, Gkritza K (2018) Income and spatial distributional effects of a congestion tax: a hypothetical case of Oregon. Transp Policy 71:28–35
16. Gu ZP (2016) The study on intelligent system of congestion pricing base on mileage. Nanjing University of Information Science & Technology, Nanjing
17. Krol R (2016) Tolling the freeway: congestion pricing and the economics of managing traffic. Mercatus Research, Mercatus Center at George Mason University, Arlington, VA
18. Eliasson J (2016) Is congestion pricing fair? Consumer and citizen perspectives on equity effects. Transp Policy 52:1–15
19. Liu ZL, Wang MJ (2011) Job accessibility and its impacts on commuting time of urban residents in Beijing: from a spatial mismatch perspective. Acta Geogr Sin 66(04):457–467
20. Liu BH, Shen FB (2011) A study on the spatial characteristics of jobs-housing structure in Beijing. Hum Geogr 26(04):40–47
21. Sommers PM (1986) Measurement of tax progressivity: a further application. Atlantic Econ J 14(1):59–62
22. Dai PS (2014) Improving measurement of tax progressivity. J Quant Tech Econ 31(02):148–160
23. Dai PS (2013) An extended study on the extended Gini index and its application. Stat Res 30(09):69–78
24. Meng B, Yu HL, Zheng LM (2012) The analysis of commuting behavior in the huge residential districts: a case study of Wangjing and Tiantongyuan in Beijing. Geogr Res 31(11):2069–2079

Social Network Analysis and Connection Strength Evaluation of Urban Tourist Attractions Using Car-Hailing Data: A Case Study of Beijing

Shixia Ma, Xuedong Yan, Xiaobing Liu and Deqi Chen

Abstract For large cities with huge tourism market, tourism travels can have significant impact on the urban traffic, which, cannot be ignored in both urban transportation and tourism planning. This paper first proposes a simple and effective spatial matching method to identify tourist travel patterns based on massive online car-hailing data. Then we construct the tourist attractions network based on the tourist movement by car. From the perspective of social network analysis, the development status of the holistic tourist attractions network and the influence of the attractions are evaluated. Finally, connection strength based Jenks natural breaks classification method is employed to divide the attractions into four levels: unconnected, weak, moderate and strong connection. Taking Beijing as case study, the main factors that affect the connection strength among attractions are the popularity of the attraction and the spatial proximity of the attraction. These findings in tourist movement can facilitate authorities and planners to develop tourism destinations and manage tourism traffic better.

Keywords Tourist attractions · Social network analysis · Connection strength · Tourist trip patterns · Car-hailing

1 Introduction

With the increase of people tourist traveling, massive tourist data can play a more important role for both urban transport planning and tourism spots management. Cities with huge tourism demand like Beijing, tourist travel patterns and travel destinations are important components of urban travel [1], but little attention is paid to this aspect. There is a difference between tourism transportation demand and urban commuting demand. When tourism transportation demand is neglected, it will lead

S. Ma · X. Yan (✉) · X. Liu · D. Chen
MOT Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, People's Republic of China
e-mail: xdyan@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_90

to the dislocation of tourism transportation demand and urban transportation supply space, and the utilization of tourism resources is not balanced. Traditional travel surveys focus on urban residents. Tourist surveys are also concentrated on the destination choice [2], as well as tourist behavior analysis in particular tourist attraction. Less research focuses on the tourist movement patterns among attractions.

With the rapid development of information technology, massive data is recorded and stored in a structured or unstructured form, making tourism research enter the era of big data. Far more existing research study tourist behavior based on diverse data, including spatial behavior [3], temporal behavior [4] and spatial-temporal behavior [5]. Travel recommendation is another hot topic. Yoon et al. [6] presented an effective recommendation based on GPS trajectories. Zheng et al. [7] predicted the next destination of individual tourists using the GPS tracking data. Tourism destination has five dimensions-spatial, temporal, compositional, social and dynamic [8]. In the destination management, tourist movements and flows act as one of the significant issues to form attraction networks in a destination [9]. Raun et al. [8] applied roaming data to analyze tourist flows, and proposed a model to forecast the number of hotel nonresident registrations. As the potential importance of understanding attraction networks in a destination, some studied attempted to look into tourist attractions network as informed by tourist movement [10] identified the spatial structure of the tourist attraction system in Seoul, South Korea based on social network analysis techniques with spatial statistics. Similarly [11] applies the Quadratic Assignment Procedure of SNA to test the relationships between region proximity, grade proximity, and tenure proximity, and the attraction network Juan and Hernández [12] reveals attractions cluster by segmenting this network using network analysis tools. In addition, limited by the data, on the mining of connection in tourist attraction visits, existing research mainly considers association rules [13]. Compared with the existing literature, this paper evaluates the connection strength among attraction and analyzes the factors affecting the relationship among attractions from the perspective of tourist movement.

The aim of this study is to explore tourist travel behavior from massive car-hailing data. Firstly, a spatial matching method is presented to identify and collect tourist trip, which shows the possibility of using the floating car data to explore the behavior of tourists. Then, the attractions network is constructed by tourist movement. From the perspective of the overall network and network nodes of social network analysis, we evaluate the development of the attraction network and the status of the attractions in the network. Finally, a simple method is proposed to evaluate the strength of the connection among the attractions.

2 Study Area and Data Preprocess

2.1 Beijing Tourist Attractions

Beijing is the capital of China, the national political center, cultural center, international exchange center, and science and technology innovation center. It is one of the first batch of national historical and cultural cities with rich tourism resources, attracting a large number of domestic and overseas tourists. According to the latest statistics of the Beijing Tourism Public Information and Consultation Platform, there are 242 A-level tourist attractions in Beijing. In this paper, taking into account the activity of the attraction, select three levels of attractions 5A, 4A, part of the 3A, a total of 98 attractions for research. The overview of the study area and attractions is shown in Fig. 1. This paper studies the behavior of tourists visit attractions during the June 2017 and National Day holiday. This period includes working days, weekends and holidays, and the source of tourists is more diverse, including local residents and foreign tourists, more comprehensively reflecting the spatial distribution structure of attractions.

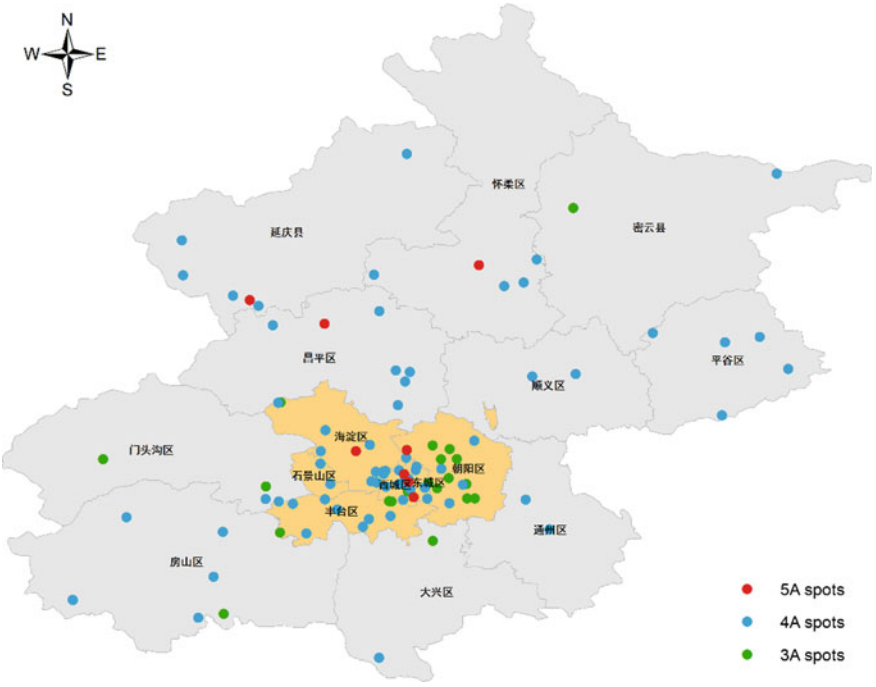


Fig. 1 The overview of the study area and attractions

2.2 Data Description

In this paper, two datasets are established. The first one is the tourist attractions information in Beijing, which includes the coordinates of the attraction and its entrance. The coordinates are obtained from the Baidu API open platform. The other is express ride orders (from 2017/06/01 to 2017/06/30 and from 2017/10/01 to 2017/10/08) in Beijing. Each express ride order includes eleven variables such as: the order ID, passenger ID, departure time, arrival time, departure latitude and longitude, arrival latitude and longitude, travel time, travel distance, and average travel speed. The passenger ID has character of uniqueness, which would be applied to distinguish different passengers and it has been data desensitization processing.

In the process of data collection and storage, the errors are caused by signal blocking or equipment failure and so on. To ensure the quality of the research data, the following express ride orders are viewed as outliers and removed from datasets:

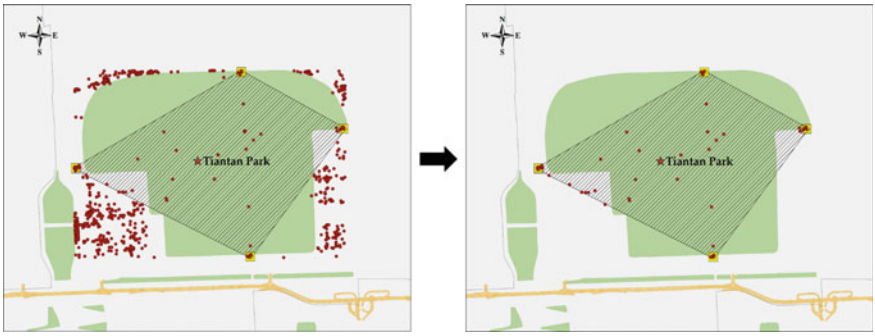
(1) Orders with Departure time before arrival time;
(2) Orders with Longitude or latitude is 0;
(3) Orders with travel distance larger than 100 km or smaller than 0.01 km;
(4) Orders where with travel time longer than 2 h or shorter than 1 min;
(5) Orders with an average speed over 100 km/h or below 1 km/h;
(6) Orders with route circuity less than 1. (The route circuity means the ratio between the actual travel distance and the Euclidian distance, and the Euclidean distance is calculated by the coordinates of the origin and destination).

2.3 Tourist Trips Identification

In the navigation map, the tourist attraction is a point. When we set the attraction as the destination, the navigation system may choose its entrance or the interior as the destination. Therefore, destinations of the order to the attraction are concentrated in the attraction or its entrance. Meanwhile, there is an intersect tool in the Arcgis10.2 overlay toolset. The intersect tool calculates the geometric intersection of any number of feature classes and feature layers, extracting the parts of the two layers that intersect the spatial relationships. The inputs can be any combination of geometry types (point, multipoint, line, or polygon). The output geometry type can only be of the same geometry or a geometry of lower dimension as the input feature class with the lowest dimension geometry (point = 0 dimension, line = 1 dimension, and poly = 2 dimension). As shown in Fig. 2, the inputs are a shadow polygon and red points, and the outputs is points within the shadow polygon.

According to the above analysis, a spatial matching method to identify tourist trips is proposed, the specific steps are as follows:

Step 1: Generate a minimum geometry polygon from the attraction and the entrance to represent the attraction area. As shown in Fig. 2, The Tiantan Park is



(a) INPUT (b) OUTPUT

Fig. 2 Diagram of intersect

taken as an example. The five-pointed star is the location of attraction. The yellow point is the four main entrances of the Tiantan Park. The shadow is the minimum geometry polygon that represents the Tiantan Park. Although the polygon and the attraction area are not completely identical, it can be seen that the orders for the Tiantan Park as the origin or destination is in this polygon.

Step 2: The order destinations constitute a point layer, and intersect with the attraction polygon layer to get tourist orders. The output orders are labeled with attraction ID. Attraction ID represents which attraction belongs to.

Step 3: In the short-term, tourists will not visit the same attraction multiple times, and the trip may be generated by the tourist attraction staff or nearby residents. It is not a tourist trip, so it is excluded from the dataset.

The trip extraction step for leaving the attraction is the same except that the point feature in step 2 is the origin. Finally, we get two trip datasets, one is the trip to the attraction, and the other is the trip from the attraction.

3 Method

3.1 Social Network Analysis

A tourist attraction system is defined as an empirical connection of tourist, nucleus, and marker [14]. The movement of tourists among different attractions represents the connectivity among attractions [12]. In this study, if a passenger has both an attraction A order and an attraction B order, or an order from attraction A to attraction B, it can believe that the passenger visited both attraction A and attraction B. Accordingly, a tourist attraction connection matrix was constructed by the amount of tourists among attractions. As the visit sequence is not considered in the study, our matrix is symmetric.

Social Network Analysis (SNA) is the process of investigating social structures through the use of networks and graph theory [15]. If the tourist attraction is regarded as a node, the connection among the attractions is the tie, SNA can be used to analyze the characteristics of the tourist attraction network, reveal the role and status of the attractions in the network, and compare the influence of different tourist attractions. This process can be carried out by UCINET6.0 and visualized in Netdraw. Since in the social network analysis, the matrix is required to be a binary matrix, we transformed the connection matrix to a binary matrix using the mean as the cut-off value. The evaluation is carried out from two perspectives: the overall network and the network node.

3.1.1 Overall Network

Network density refers to the proportion of direct ties in a network relative to the total number possible [16], which measures the overall tightness of the network. The higher the density in a network, the more connections, the closer it is to 1, indicating that the more frequent the flow of tourists among the various attractions. The network centralizations characterize the difference in various nodes in the network and describe the consistency of the overall network. Examples of network centralizations commonly include degree centralization and betweeness centralization. The degree centralization is gathering trend of the nodes in the network. If the degree centralization is larger, the greater the possibility that the flow of tourists will gather or spread around the attraction with high degree. The betweeness centralization is the quantification of the betweeness centrality of the node. The higher the value, the more nodes in the network may be divided into multiple small groups and rely on one node to transfer the relationship. In our network, it means that there may be clustering in the tourist attraction system. The purpose of Core-Periphery structural analysis is to study which nodes in the social network are at the core and which are at the edge. Through the Core-Periphery, we can judge the hierarchical structure of the attraction and analyze the core attractions that mainly attract tourists.

3.1.2 Network Node Analysis

Centrality is one of the focuses of SNA, referring to a group of metrics that aim to quantify the “importance” or “influence” of a particular node within a network [17]. According to different calculation methods, there are three common types of centrality: degree centrality, betweeness centrality, and closeness centrality. The degree centrality is the number of immediate connections an attraction has in a network, which can be used to measure an attraction’s level of involvement, prestige, or dependence in the network [10]. The betweeness centrality refers to the times of a node acts as the shortest bridge between the other two nodes. The betweeness centrality measures the ability of a node as an intermediary. An attraction with high betweeness centrality have strong control over other attractions. The closeness centrality

is meaningful when all nodes in the network are connected. In this study, there are isolated nodes in the network. Thus, closeness centrality was not computed. The degree centrality can be formulated as follows:

Degree centrality, for attraction A_i:

$$C_D(\mathrm{A}_i) = \sum_{j=1}^{n} x_{ij}(\mathrm{i} \neq \mathrm{j}) \tag{1}$$

where, x_{ij} is the value of the tie from attraction i to attraction j (the value being either 0 or 1). n is the number of attractions in the network.

3.2 Connection Strength Evaluation

In social network analysis, the connection among attractions are binary. The value is 1 when two attractions have common tourists and there is a tie between them, otherwise, the value is 0. We know which attractions have great influence, but we don't know the close connection pairs, and we don't know the strength of the connection among the attractions. Therefore, to further reveal the connection strength among attractions, the Jenks natural breaks classification method (JNBC) is employed. It is one of the most common methods for data classification. JNBC seeks to reduce the variance within classes and maximize the variance between classes [18]. Using the Jenkspy to calculate the natural breaks of the connection value among the attractions, according to which the connection between the two attractions can be divided into four levels: unconnected, weak, moderate and strong connection, labeled by 0, 1, 2 And 3 respectively. Thus, the connection strength can be formulated as follows:

We define a variable $S(\mathrm{C}_{ij})$ to represent connection strength index between attraction A_i and attraction A_j, as shown in Eq. (2), Where b_0, b_1, b_2, b_3 are natural breaks of connection value, sorted in ascending. c_{ij} is connection value.

$$S(\mathrm{C}_{ij}) = \begin{cases} 0, \mathrm{c}_{ij} = 0 \\ 1, \mathrm{c}_{ij} \in [b_0, b_1) \\ 2, \mathrm{c}_{ij} \in [b_1, b_2) \\ 3, \mathrm{c}_{ij} \in [b_2, b_3] \end{cases} \tag{2}$$

As illustrated in Fig. 3, for example, a simple case is shown in four tourist attractions (labeled A1, A2, A3 and A4). In the connection matrix (see Fig. 3a), the rows and columns index are attractions and the cell value refer to the number of tourists among attractions. Then, we transform the connection matrix into a connection strength matrix (see Fig. 3b) by Jenks. After obtaining the connection strength matrix, the Arcgis10.2 is applied for visual display to reflect the connection strength distribution among the attractions.

	A1	A2	A3	A4
A1	-	0	25	5
A2	0	-	172	53
A3	25	172	-	33
A4	5	53	33	-

(a) Connection matrix

	A1	A2	A3	A4
A1	-	0	2	1
A2	0	-	3	2
A3	2	3	-	2
A4	1	2	2	-

(b) Connection strength matrix

Fig. 3 Data matrix

4 Results

4.1 Tourist Attractions Network Profiles

Applying the proposed method of tourist trip identification, orders related to 63 attractions was extracted, accounting for 64.3% of the given tourist attractions. Some of the remaining attractions are missing information, while others are far from the city center where there are almost no tourists taking taxis. The travel data of attractions include seven 5A attractions, forty-two 4A attractions, thirteen 3A attractions and Lama Temple which is not an A-level attraction, but it is very active. These travel datasets cover almost all the popular attractions in Beijing, which are representative and can be used to study the tourist movement characteristics among tourist attractions.

The density of Beijing attractions network is 21.44%, indicating that the tourist movement network is relatively loose, and the connection among the attractions are not tight enough. The degree centralization index is 25.15%, which is relatively high. The network has a concentrated trend, and the central attractions have obvious influence. The betweeness centralization is 3.07%, which is relatively small, indicating that the interdependence among the attractions is relatively low and most of attractions are directly connected. Through the analysis of the overall network, it is noticed that the overall connection among Beijing tourist attractions is weak, the influence of the attractions varies greatly, and tourists may focus on several important attractions.

After calculating the core degree of 98 attractions, Core-Periphery module divides them into core and edge levels, of which 22 attractions are core nodes: Summer Palace, Tiantan Park, Forbidden city, Chaoyang Park, Ditan Park, Jingshan Park, Beijing Zoo, etc., and other attractions are the edges. The density of the core attraction is 66.2%, which is much higher than the overall network density by 21.44%, indicating that the core attractions are closely connected. The network density of the edge attractions is only 0.4%, which is far lower than the overall network density. The relationship among the edge attractions is extremely weak, the connections among

the attractions are not tight and even some attractions are not connected to each other. The polarization of network density in the core and edge attractions reveals that the network structure of Beijing tourist attractions is extremely uneven. The density among the core attractions and the edge attraction is 5.1%, showing that the core attraction and the edge attraction are not closely connected, and the core attractions hardly promote the development of the edge attractions. Therefore, while maintaining the good development of the core attractions, on the one hand, it is necessary to enhance the competitiveness of the edge attractions, on the other hand, it is vital to strengthen the cooperation between the core attraction and the edge attraction. The core attractions drive the development of the edge attractions, thus promoting the healthy development of the city tourism system.

4.2 Tourist Attractions Profile

The centrality of tourist attractions were calculated using UCINET6.0 and top 10 are shown in Table 1. Employing the Netdraw, we provide a visual network diagram in Fig. 4 based on the results of degree centrality and attraction grade. In this figure, the isolated node is removed. The average value of degree centrality is 5.102, which implied that each attraction is might be directly connected to five attractions. The top 10 attractions in degree centrality are at central in the attraction system, with strong agglomeration and radiation function, and are concentrated destinations for tourists. As shown in Table 1 and Fig. 4, attractions with high degree centrality are a popular attractions that we are familiar with, and high-level attractions may have high degree centrality. In addition to the Forbidden City, the Summer Palace and the Tiantan Park, the Jingshan Park and the Beihai Park are also important nodes because they are close to the Forbidden City, attracting a large number of tourists. An interesting

Table 1 Centrality of tourist attractions

Rank	Attractions	Degree	Rank	Attractions	Betweeness
1	Tiantan Park	29	1	Tiantan Park	151.10
2	Beijing Zoo	28	2	Chaoyang Park	142.56
3	Chaoyang Park	27	3	Beijing Zoo	65.72
4	Beihai Park	26	4	Xinglong Park	65.15
5	Summer Palace	25	5	Summer Palace	57.34
6	Forbidden City	25	6	Jingshan Park	55.17
7	Jingshan Park	21	7	Beihai Park	50.24
8	Lama Temple	21	8	Forbidden City	38.55
9	Taoranting Park	17	9	Ritan Park	31.98
10	Ritan Park	17	10	Lama Temple	24.99

Note degree: degree centrality; betweeness: betweeness centrality

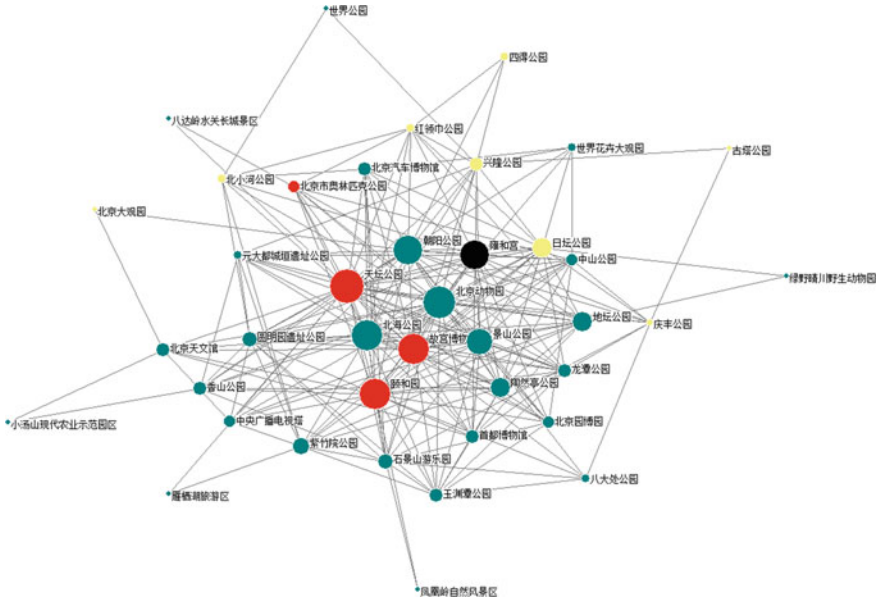


Fig. 4 Tourist network of Beijing

finding is that the Lama Temple is a non-A-level attraction, but an important node. The Lama Temple is the highest-ranking Buddhist temple in the middle and late Qing Dynasty, where attracts lots of people to pray.

The betweeness centrality refers to the ability of the attraction to act as an intermediary in the tourist movement network, which can reflect the control ability of the node to tourists flow in the network to some extent. The Tiantan Park and the Chaoyang Park are not only at the central position, but also control the connection among other attractions and play an important role in the attractions system. However, Xinglong Park is a node with low centrality but high centrality, indicating that it controls the connection with some relative edge nodes, plays an important role in the local, and improves the overall connection of the network.

It should be noted that the size of nodes depends on the value of degree centrality. Nodes in red refer to 5A attractions; nodes in green refer to 4A attractions; nodes in yellow refer to 3A attractions; node in black refers to The Lama Temple.

4.3 Connection Strength Distribution

The connection strength of the attractions are shown in Fig. 5. There are 5 pairs of strong connections, 11 pairs of moderate connections, and 196 pairs of weak connections. From Table 2, it shows 16 attraction pairs with strong or moderate connection. The Forbidden City, the Summer Palace and the Tiantan Park is strongly connected

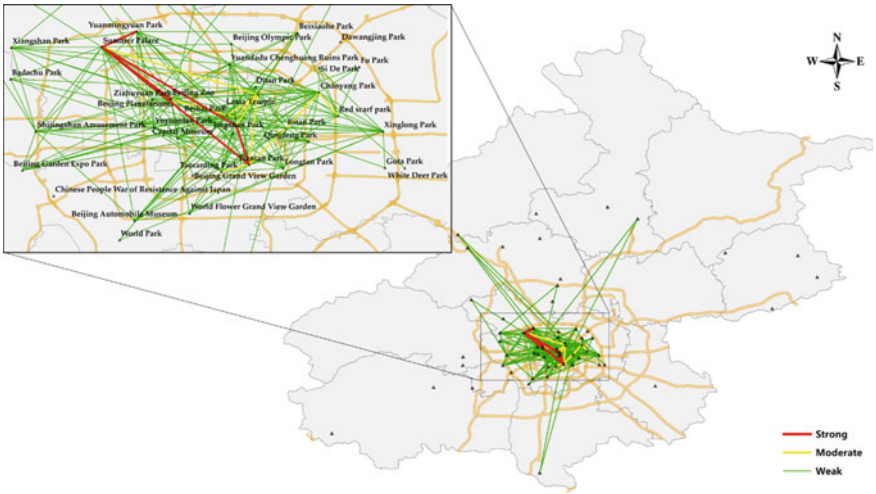


Fig. 5 Distribution of connection strength

Table 2 Connection strength of tourist attractions

Connection strength	Attraction pairs	Connection strength	Attraction pairs
Strong	Summer Palace—Forbidden City	Moderate	Lama Temple—Forbidden City
	Summer Palace—Tiantan Park		Lama Temple—Summer Palace
	Summer Palace—Yuanmingyuan Park		Lama Temple—Beihai
	Summer Palace—Beijing Zoo		Lama Temple—Jingshan
	Forbidden City—Tiantan Park		Beihai Park—Jingshan Park
Moderate	Tiantan Park—Taoranting		Beihai Park—Beijing Zoo
	Tiantan Park—Ritan		Chaoyang Park—Ritan Park
	Tiantan Park—Lama Temple		Chaoyang Park—Red Scarf Park

to each other. The Lama Temple is strongly connected with the Forbidden City, the Summer Palace, and the Tiantan Park. In addition to the strong connection among the popular attractions, the connection among attractions with close geographical location is also strong, forming a tour route of the group. For example, the Summer

Palace and Yuanmingyuan Park; Tiantan Park, Taoranting and Ritan Park; Beihai, Jingshan Park and Beijing Zoo; Chaoyang Park, Ritan and Red Scarf Park.

5 Discussion and Conclusion

The contribution of this study can be summarized in three aspects. First, although taxi trip data is widely used for resident travel research, especially commuting, there is little attention on other travel purposes, such as leisure. This study proposes a method for spatial matching using Arcgis10.2 to identify tourists travel from taxi travel data, making up for the lack of data in traditional tourism research. The travel data obtained from the GPS of the vehicle has a long period and a large spatial scale, which can provide data support for more travel research. The method proposed in this study provides new insight for taxi travel data mining, especially for tourism travel.

Second, the tourist attractions network is very important in the destination management, and the increasing research begins to focus on demand-driven network relationships among attractions in a destination [11]. This study builds a tourist attractions network based on tourists movement, and evaluates the network from the perspectives of overall network and network nodes. New knowledge was contributed to the field of tourism network. In the analysis of the overall structure of the network, four indicators of density, degree centralization, betweeness centralization and Core-Periphery are applied. The results show that the overall structure of the tourist movement network in Beijing is loose and unevenly developed, and there is great potential for developing tourist routes. The core attractions are closely connected to each other, but the edge attractions are not only weakly attractive to tourists, but also not closely connected to the core attractions. In the analysis of network nodes, the influence of the attractions are compared and analyzed with two indicators: the degree centrality and the betweeness centrality. There are many tourist attractions in Beijing, but tourists are only interested in some attractions. These attractions are traditional attractions with high reputation and brand awareness. They are at the center of the network. These attractions also have strong intermediary capabilities and control the flow of tourists in other attractions in the network.

Third, the analysis of the connection among the attractions can provide support for traffic management among the attractions to meet the tourist traffic demand during the holiday. Based on the amount of tourists among attractions, this study proposes a method to divide the connection strength among the attractions, identify the strong connection pairs of attractions, and try to analyze the factors affecting the strength of the relationship among the attractions. According to the results, it is found that an attraction with strong or moderate connection strength is almost a central attraction in the network. In addition, when the central attraction has an enough influence on the nearby attraction, there will be a strong connection between them, such as the Yuanmingyuan Park and Red Scarf Park, which are closely connected with the Summer Palace and the Chaoyang Park respectively. Based on this, we speculate on

two main factors that cause the strength of the connection among the attractions: one is the attraction of the tourist attractions to tourists and its influence in the network; the second is the proximity of geospatial location.

6 Suggestion and Future Work

Based on the analysis of the tourist attraction network and relationship among attractions, the problems of tourism existing in tourism management and traffic planning are explored, and some suggestions with practical application value can be provided for managers. Beijing has abundant tourism resources, but its development is uneven. It is necessary to strengthen the overall network connection and establish a good competition and cooperation system to enable the core attractions to drive the development of the edge attractions. There are many tourists in the core attraction, which leads to the pressure of reception in the attraction. However, there may be almost no tourists in the surrounding attractions. Thus, it is necessary to enhance the publicity and popularity of edge attractions. At the same time, form some regional tourist routes, evacuate the tourists gathered in the popular attractions, increase the tourists in the surrounding attractions, and enhance the tourist experience.

Traffic is the basic guarantee for the tourist movement among attractions. For traffic managers, on the one hand, it is imperative to strengthen the transportation infrastructure among the attractions with high tourist flow, to meet the demand for tourism travel, and to do the traffic diversion work during the holidays to ensure the smooth flow of the urban transportation system. On the other hand, it is useful to depend on traffic to guide the tourist flow, to provide transportation services for isolated nodes or edge attractions, and to enhance their centrality in the tourism network.

Although this study explored the attraction network and puts forward some insights, it is not deep enough and comprehensive. Future research should consider more factors affecting the attraction, such as basic service facilities, traffic accessibility, etc.

Acknowledgements This work is financially supported by National Natural Science Foundation (91746201, 71621001). Many thanks to the Institute of Policy Studies, DiDi Company for providing the data in this study.

Conflicts of Interest The authors declare that there is no conflict of interest in any aspect of the data collection, analysis, or the funding received regarding the publication of this paper.

References

1. Hasnat MM, Hasan S (2018) Identifying tourists and analyzing spatial patterns of their destinations from location-based social media data. Transp Res Part C 96:38–54

2. Yu P, Xiao L, Yan-Yang MA et al (2016) The study on the factors influencing the citizen elderly' choice for tourism destination—based on a market survey of middle-aged and elderly people in Beijing Chaoyang. Sci Technol Ind
3. East D, Osborne P, Kemp S, Woodfine T (2017) Combining GPS & survey data improves understanding of visitor behaviour. Tour Manag 61:307–320
4. Edwards D, Griffin T (2013) Understanding tourists' spatial behaviour: GPS tracking as an aid to sustainable destination management. J Sustain Tourism 21(4):580–595
5. Shoval N, McKercher B, Birenboim A, Ng E (2015) The application of a sequence alignment method to the creation of typologies of tourist activity in time and space. Environ Plan 42(1):76–94
6. Yoon H, Zheng Y, Xie X, Woo W (2010) Smart itinerary recommendation based on user-generated GPS trajectories. In: Proceedings of the 7th international conference on ubiquitous intelligence and computing, Xi'an, China
7. Zheng W, Huang X, Li Y (2017) Understanding the tourist mobility using GPS: Where is the next place? Tour Manag 59:267–280
8. Raun J, Ahas R, Tiru M (2016) Measuring tourism destinations using mobile tracking data. Tour Manag 57:202–212
9. Lew A, McKercher B (2006) Modeling tourist movements: a local destination analysis. Ann Tourism Res 33(2):403–423
10. Kang S, Lee G, Kim J, Park D (2018) Identifying the spatial structure of the tourist attraction system in South Korea using GIS and network analysis: An application of anchor-point theory. J Destination Mark Manag 9:358–370
11. Liu B, Huang S, Fu H (2017) An application of network analysis on tourist attractions: the case of Xinjiang, China. Tourism Manag 58:132–141
12. Hernández JM, Kirilenko AP, Stepchenkova S (2018) Network approach to tourist segmentation via user generated content. Ann Tourism Res 73:35–47
13. Versichele M, de Groote L, Bouuaert MC, Neutens T, Moerman I, Van de Weghe N (2014) Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: a case study of Ghent. Belgium Tourism Manag 44:67–81
14. Leiper N (1990) Annals of tourism research. Tourist Attraction Syst 17(3):367–384
15. Otte E, Rousseau R (2002) Social network analysis: a powerful strategy, also for the information sciences. J Inf Sci 28(6):441–453
16. Xu G et al (2010) Web mining and social networking: techniques and applications. Springer, Berlin, p 25. ISBN 978-1-4419-7734-2
17. Liu B (2011) Web data mining: exploring hyperlinks, contents, and usage data. Springer, Berlin, p 271. ISBN 978-3-642-19459-7
18. Jenks GF (1967) The data model concept in statistical mapping. Int Yearbook Cartogr 7:186–190

Visualization of Spatio-Temporal Traffic Performance in Urban Road Network Based on Grid Model

Liwei Wang, Yingnan Yan and Deqi Chen

Abstract Traffic congestion monitoring is a very important problem in big cities. This paper describes a novel computation way of traffic performance index in urban road network. Floating car data are used in the extraction of traffic attributes. In addition, we use the grid model to reconstruct the road network and present the visualization approach of spatio-temporal traffic performance based on it. We take the 5th ring area of Beijing as a case area. The case results indicate that traffic performance is worse in the PM than in the AM. Moreover, it is worse in the north and middle area than in the south area. This method efficiently distinguish the congestion's area and time of the urban road network. It is much helpful with traffic managers in city traffic management.

Keywords Floating car data · Delay · Speed · Index · Traffic management

1 Introduction

Traffic congestion has become a very important problem in big cities. In 2019, Beijing's expenditure budget for transportation management projects is about 1.4 billion yuan. Beijing Traffic Management Bureau [1] Spending much money, traffic managers are willing to identify when and where the traffic congestions occurred. However, there is lack of the efficient way to visualize the traffic performance in the whole city area. By the development of the technology of floating car data (FCD), this work has new methods to deal with. Brockfeld et al. [2] had pointed that the FCD system can particularly detect congestion condition by the experiment of Germany

L. Wang (✉) · D. Chen
MOT Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, People's Republic of China
e-mail: wangliwei@bjtu.edu.cn

Y. Yan
The High School, Beijing Jiaotong University, Beijing 100044, People's Republic of China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_91

FCD. In this paper, we give a novel method to work with the visualization of spatio-temporal traffic performance in urban road network based on floating car data.

The traffic performance has many identifications. In Highway Capacity Manual [7], the traffic performance of road segment is classified by the travel speed and the traffic performance of intersection is classified by the control delay. De Fabritiis et al. [3] gave an estimation model of traffic speed based on FCD and made an experiment in the Rome ring road. Kong et al. [10] used a fuzzy comprehensive evaluation method to identify the traffic performance based on the traffic speed, density and volume. Li et al. [11] used five index parameters including the degree of saturation, the average stopped delay, the queue length, the conflict ratio, and the separation ratio to evaluate the performance of signalized urban intersections under mixed traffic conditions by gray system theory. However, the road network is more complicated, which owns both urban roads segments and intersections. Some researches [5, 9] used FCD to analyze the traffic congestion in urban road network based on the speed performance and travel time index. Enlightened from above papers, we choose the speed ratio, delay ratio as the traffic performance indicators.

The grid model is a good method to integrate the FCD with calculation and traffic evaluation. It was well used to simulate the Manhattan road network with the intersections of roads [15]. In addition, it is helpful with the spatio-temporal filtering of speeds and the estimation of travel time. Van Lint [16] and Demiryurek et al. [4] used the grid model to generate the spatio-temporal freeway network of LA County in United States. It is important to determine the grid size while using the grid model to deal with the traffic problem. Recently, He et al. [6] used grid model to extract traffic states based on FCD. In this research, the road network was divided into small cells and each cell had the side length of 100 m. However, small grids have some defects with the extraction of travel time. The trajectory may cross over more than one grids and it is hard to determine the relationship between adjacent grids. Therefore, we choose the proper grid size to describe the urban road network in this paper.

In addition, there are some other models dealing with the traffic performance problem. Huang et al. [8] construct a realistic and large-scale mobility model of Shanghai urban vehicular network by FCD. Pongpaibool et al. [14] used fuzzy techniques to evaluate the road traffic congestion. Li et al. [12] evaluated the performance of vehicle-based mobile sensor networks for traffic monitoring based on the macroscopic traffic-flow theory. Moreover, Liu et al. [13] evaluated the traffic status based on fuzzy clustering and RBF neural network.

According to previous studies, the computation of traffic performance index did not reach a consensus. In addition, many of them lack a good visualization approach to represent the spatio-temporal result of the whole road network (not a road segment or an intersection). In this paper, we give the novel computation way of traffic performance index, which is easy to obtain and efficient to represent the traffic performance level. Moreover, this paper gives the visualization approach of traffic performance in both space and time dimension based on the grid model. In addition, this visualization approach is well read for everyman, including the traffic managers and the vehicle drivers.

The remainder of this paper is organized as follows. The second section describes the proposed method including the way of data preprocessing, map grid, traffic attributes extraction and the visualization method. The third section is the visualization result of traffic indicators and traffic performance with the case study of the 5th ring area in Beijing. And the final section summarizes the conclusion and recommendations.

2 The Proposed Method

2.1 Data Preprocessing

The raw data of floating car are taxi GPS trace points of Beijing in November, 2012. There are totally 12,000 taxis, generating 44.4 GB trace data in a month. The attributes contain vehicle number, taxi operation state, GPS time, longitude, latitude, vehicle speed, and GPS state. Specially, Taxi position is determined by longitude and latitude based on the Worldwide Geodetic System 1984 (WGS84) coordinate system. The taxi operation state presents whether there are passengers riding in taxis. The GPS state shows the validation of GPS device, and 0 represents that the GPS device could not work (Table 1).

The Preprocessing process contains 4 steps. Step 1 is to remove the invalid GPS data where "GPS State" is 0. Step 2 is to filter out data of taxis carrying no passengers because vacant taxis have the action of searching for passengers in low speed, which has a harmful effect on the result. Step 3 is to eliminate the data out of range, for instance, the taxi speed out of the range from 0 to 120 km/h and the position out of the research region. Step 4 is to remove the data in non-workdays because traffic performance is quite different from workdays to non-workdays, and we focus on the traffic state in workdays in this research.

Table 1 The data attributes of FCD system

Field name	Field type	Field description
Vehicle number	String	6 bytes characters, marking each vehicle
Taxi operation state	Integer	0 = no passenger; 1 = serving passenger
GPS time	Timestamp	Accurate to second
Longitude	Floating	Accurate to six decimal places
Latitude	Floating	Accurate to six decimal places
Vehicle speed	Integer	Kilometer per hour
GPS state	Boolean	1 = valid; 0 = invalid

2.2 Map Grid

Floating car trajectory data and grid method are used to extract the traffic attributes of grid, which can improve the computing efficiency by matching the data to the grid. It is important that the grid size should be neither so small or so large. If the grid size is so small, the trace data would stride across more than one grid so to make it difficult to extract the traffic state. If the grid size is so large, we could not distinguish the difference between small areas. Hence, it is better to determine the grid size by the sampling rate of floating car trajectory data and the urban road network structure.

The static attributes of grids contain two parts. One part is the geographical position attributes including the position of grid, coordinates of grid apexes and the coordinates of intersection points between roads and grid boundaries. Geographical position attributes reflect the traffic road position and the relationship between roads and grids. The other part is the topological relation between entry points and exit points in each grid. Topological relation reflects the traffic route information of urban roads, and each node pair represents each traffic route. There are many entry points and exit points on the grid boundaries, and each point reflects the relationship of adjacent grids.

In this paper, the urban road network within fifth ring area in Beijing is divided into 30 * 30 grids and each grid's border length is 1 km. Grids are numbered from southwest as 0 to northeast as 899. The distribution of grid is shown in Fig. 1.

2.3 Traffic Attributes Extraction

The traffic performance of urban roads is officially described by the speed or traffic delay [7]. Hence, we could define the two main traffic attributes that are the speed ratio and delay ratio. The speed ratio is defined by the free flow speed divided by the average speed as shown in function (1). In which T_a denotes the average travel time and T_f denotes the free flow travel time. In addition, the delay ratio is defined by the average travel time divided by the free flow travel time as shown in function (2). In which V_f denotes the free flow speed and V_a denotes the average speed. The delay ratio reflects the intuitive congestion feeling of drivers and passengers in the grid, which represents that it takes several times more time for vehicles than the free flow state from entering to leaving the grid. Moreover, the speed ratio is the ratio of the current speed to the free flow speed, which can reflect the congestion level under different road conditions.

$$r_d = T_a / T_f \tag{1}$$

$$r_v = V_f / V_a \tag{2}$$

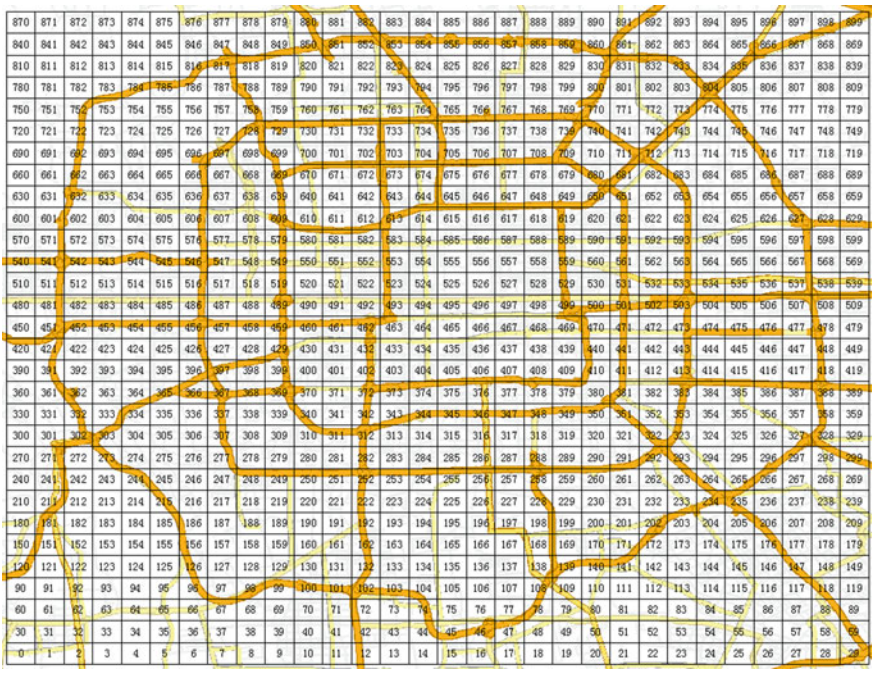


Fig. 1 Grid numbers within Beijing fifth ring area

The traffic attributes of each grid are extracted based on the grid model. The grid model defines the geographical position and the topological relationship of all the grids. Therefore, the floating car data could have the relation of one grid and each grid could aggregate the attributes of traffic state. The urban road network is divided into serval grids so each grid has the information of a simple road network. If the floating car pass though one of the grid, it must has a route of serval floating car data points with longitudinal and latitudinal coordinates. When many floating car data pass though the grid, the traffic state of each route can be described with statistical methods. Therefore, the average speed and average travel time of each grids are calculated by the floating car data of each grid. To extract the free flow speed, we need to sort the floating car travel speed of each route and choose the ten-percentile record as the free flow speed. The extraction method of free flow travel time is the same as the speed.

2.4 Visualization of Traffic Performance Index

The traffic performance is determined by the combination of traffic speed ratio and delay ratio, as shown in function (3). In which r denotes the ratio indicator and n

Table 2 The color representation and description of traffic performance index

The traffic performance index	Color	Description
0–0.5	Dark green	The traffic is unobstructed
0.5–1.0	Light green	The traffic is smooth and almost without congestion
1.0–1.5	Yellow	The traffic state is normal with some congestion
1.5–2.0	Orange	The traffic has frequent congestion
>2.0	Red	The traffic is almost blocked

denotes the indicator number. In this way, the traffic performance index has a lower bound of zero which denotes the traffic performance is excellent and everywhere is free flow speed and free flow travel time. When the traffic performance index is greater than zero, the number of it indicates the total ratio of traffic conditions. The larger of the number indicates the worse of the traffic performance.

$$I = \sum_{d,v}^{n} r - n \tag{3}$$

The traffic performance index could be calculated in each grid and each period of day. Therefore, the space distribution in a specific period of traffic performance could be represented as a color map. The color representation and description of traffic performance index is shown in Table 2. The dark green has the index of 0–0.5 interval and it denotes that the traffic state is unobstructed. And the light green color has index of 0.5–1.0 interval which denotes that the traffic is smooth and almost without congestion. In addition, the yellow color means that the traffic state is normal with some congestion and has the index interval of 1.0–1.5. Moreover, the orange color with index interval of 1.5–2.0 represents that the traffic has frequent congestion. Last, the red color represents the most congestion condition of all the colors while the traffic performance index is larger than 2.0.

The color map is a well visualization method to show a traffic performance of the entire urban road network in a slice of time based on the grid model. Because it can clearly reveal the space and time information of traffic performance. Therefore, the traffic managers could take advantage of this information to control the traffic congestion. In addition, the software of ArcGIS is a suitable tool to deal with the color map works. And we use this software to draw the color figures of the distribution of traffic indicators and performance index in this study.

3 Result

3.1 Traffic Ratio Indicator Result

By the grid model, we obtain the spatio-temporal result of speed ratio and delay ratio. The 5th ring region of Beijing is chosen as a study area. The period of morning peak hour (7:00–9:00) is chosen for the representation of traffic ratio indicator space distribution. Colors in the distribution map denotes the value of the ratio indicators as shown in Figs. 2 and 4 (The value range of each color is chosen artificially in order to clearly show the change of hierarchy). When the value increases, the colors change from green to yellow to red, also the traffic performance becomes worse. The time series of indicators are shown in Figs. 3 and 4. The time is divided into 144 pieces and each piece of time is 10 min. And the value of it is the average value of ratio indicator in the whole study area.

From the space distribution result of delay ratio as shown in Fig. 2, we find that the delay ratio is mainly larger (greater than 3.0) in the Lianhuachi East Road area, the southwest of the 2nd ring expressway area, the southeast, west and north of the 3rd ring expressway area, the airport expressway area and the Gaobeidian area. In addition, the delay ratio is lower (smaller than 1.5) mainly far from the center of

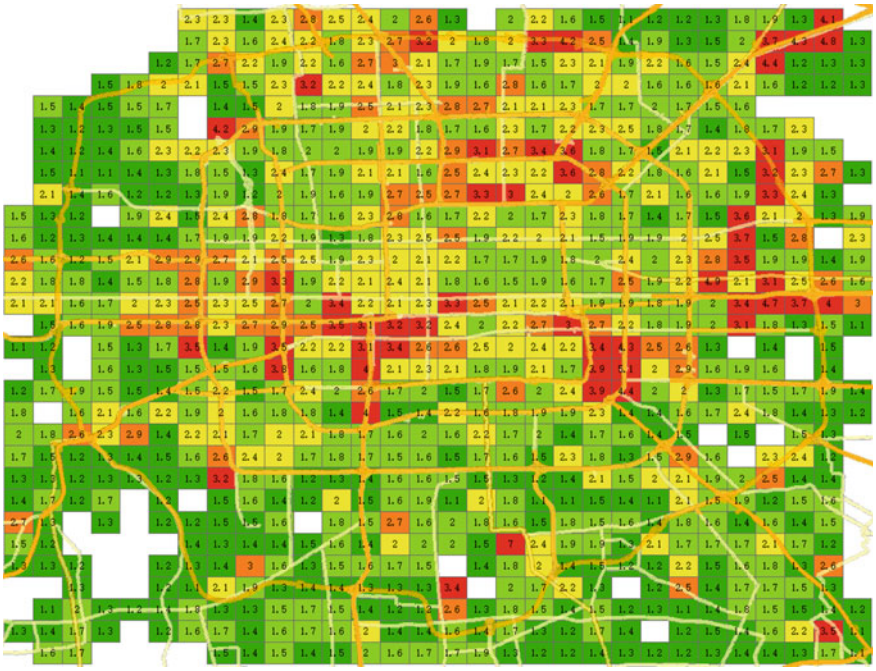


Fig. 2 The space distribution of delay ratio in morning peak hour

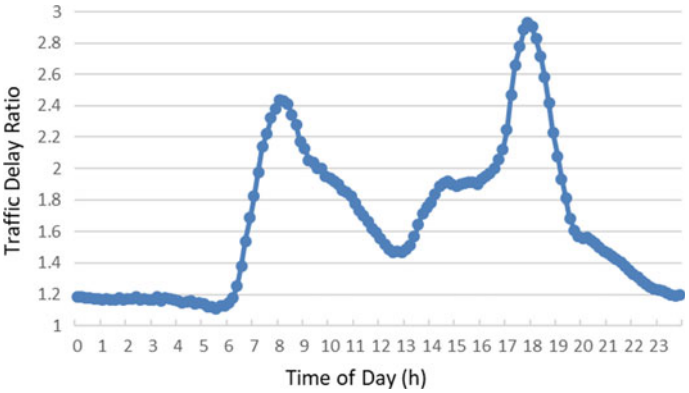


Fig. 3 The time series of delay ratio

Fig. 4 The space distribution of speed ratio in morning peak hour

Beijing, especially in the outside of the 5th ring region. The time series of delay ratio is shown in Fig. 3. The traffic delay ratio starts to jump at 6:00 and ends increasing at 8:00. The first peak of traffic delay ratio is 2.44 at 8:00. Then, it decreases to 1.47 at 12:50. The suddenly increasing of indicator in PM is at 17:00 and have a peak value of 2.93 at 17:50. Then it decreases until the next morning.

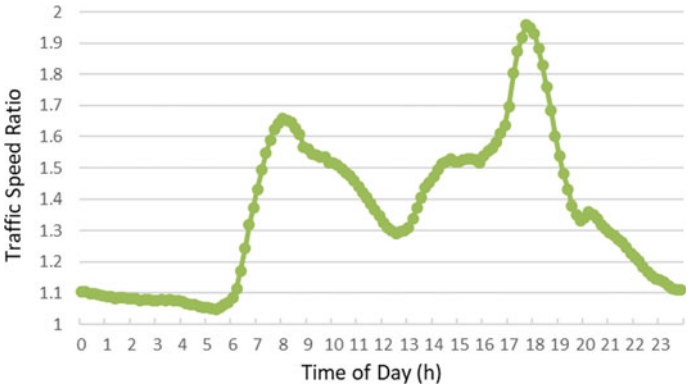


Fig. 5 The time series of speed ratio

The spatio-temporal result of speed ratio is shown in Figs. 4 and 5. The larger value (larger than 2.2) of speed ratio is mainly distributed in the north, west and east of the 2nd ring, the 3rd ring area and other dispersed grids. And the small value (small than 1.3) of speed ratio is mainly distributed at the edge of Beijing. Moreover, the time trend of speed ratio is very similar to the trend of delay ratio. The peak and trough time are the same. The only difference is the value size. In which the speed ratio is 1.66 at AM peak and 1.96 at PM peak.

3.2 Traffic Performance Index Result

The spatio-temporal result of traffic performance index is obtained as shown in Figs. 6 and 7. Colors indicate the interval of index by Table 2. The worse traffic performance (when the index is larger than 1.5) of morning peak hour is mainly distributed in the north and the middle area compared with the south area. There is a large congestion area in the neighborhood of Muxidi and Wukesong metro Station and extends to the Qianmendong Street. The time series of performance index is similar to the trend of traffic indicators. The index reaches 2.10 of AM peak hour at 8:00 and 2.88 of PM peak hour at 17:50. Therefore, the traffic performance is worse in the PM than in the AM hours. From Fig. 7, we can also identify the morning and evening rush hour. The morning rush hour is approximately from 7:20–9:40 (when the average index is larger than 1.5), and the evening rush hour is approximately from 16:10–19:00 (when the average index is larger than 1.5). Meanwhile, the relatively unobstructed moments of daytime is about 0:00–6:40, 11:40–13:20 and 17:40–24:00 (when the average index is less than 1.0).

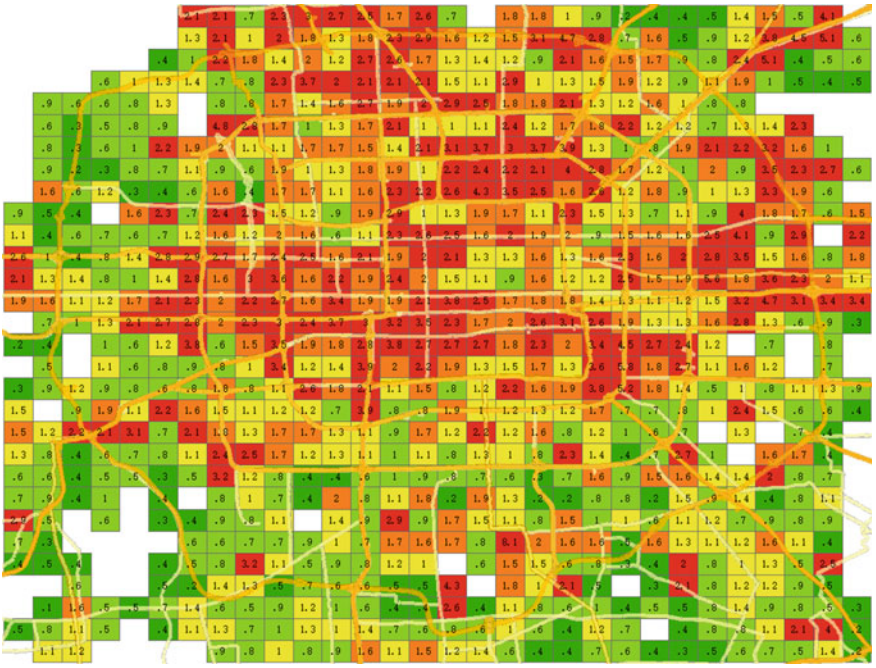


Fig. 6 The space distribution of performance index in morning peak hour

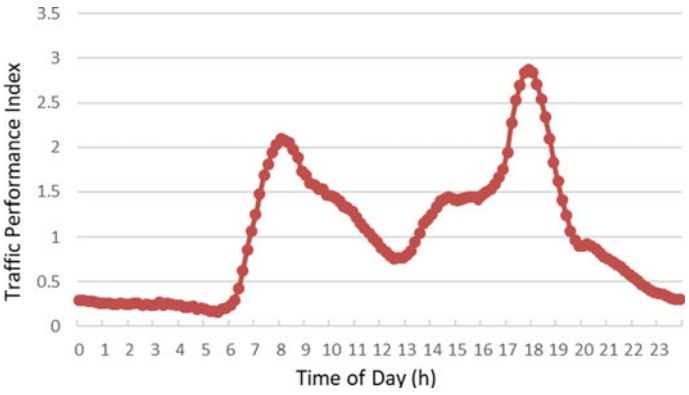


Fig. 7 The time series of performance index

4 Conclusion

In this paper, we give a visualization method of spatio-temporal traffic performance in urban road network based on the floating car data and grid model. The traffic attributes including the traffic delay ratio and speed ratio are extracted as indicators.

We define the traffic performance index and give the visualization approach of it in the grids. From the result, we obtain the space distribution of AM peak hours and time series of traffic delay ratio, speed ratio and traffic performance index. The traffic performance is worse in the PM than in the AM. Moreover, it is worse in the north and middle area than in the south area. The index reaches 2.10 of AM peak hour at 8:00 and 2.88 of PM peak hour at 17:50. In addition, we redefine the morning and evening rush hour. The morning rush hour is approximately from 7:20–9:40, and the evening rush hour is approximately from 16:10–19:00. Meanwhile, the relatively unobstructed moments of daytime in Beijing is about 0:00–6:40, 11:40–13:20 and 17:40–24:00. This method efficiently distinguish the congestion's area and time of the urban road network. Therefore, it is much helpful with traffic managers in city traffic management.

Acknowledgements This work is financially supported by National Natural Science Foundation (91746201, 71621001) and Postgraduate Innovation Project (2019YJS081).

References

1. Beijing Traffic Management Bureau (2019) Budget information of public security and traffic administration of Beijing Public Security Bureau. http://jtgl.beijing.gov.cn/jgj/czyjs/98326/index.html
2. Brockfeld E, Lorkowski S, Mieth P, Wagner P (2007, June) Benefits and limits of recent floating car data technology–an evaluation study. In: 11th WCTR Conference, Berkeley, USA
3. De Fabritiis C, Ragona R, Valenti G (2008, October) Traffic estimation and prediction based on real time floating car data. In: 2008 11th international IEEE conference on intelligent transportation systems. IEEE, New York, pp 197–203
4. Demiryurek U, Pan B, Banaei-Kashani F, Shahabi C (2009, November) Towards modeling the traffic data on road networks. In: Proceedings of the second international workshop on computational transportation science. ACM, New York, pp 13–18
5. He F, Yan X, Liu Y, Ma L (2016) A traffic congestion assessment method for urban road networks based on speed performance index. Proc Eng 137:425–433
6. He Z, Zheng L, Chen P, Guan W (2017) Mapping to cells: a simple method to extract traffic dynamics from probe vehicle data. Comput-Aided Civil Infrastruct Eng 32(3):252–267
7. Highway Capacity Manual (2010) Transportation Research Board of the National Academies, Washington, D.C.
8. Huang HY, Luo PE, Li M, Li D, Li X, Shu W, Wu MY (2007) Performance evaluation of SUVnet with real-time traffic data. IEEE Trans Veh Technol 56(6):3381–3396
9. Jenelius E, Koutsopoulos HN (2013) Travel time estimation for urban road networks using low frequency probe vehicle data. Transp Res Part B: Methodol 53:64–81
10. Kong X, Xu Z, Shen G, Wang J, Yang Q, Zhang B (2016) Urban traffic congestion estimation and prediction based on floating car trajectory data. Future Gener Comput Syst 61:97–107
11. Li J, Yue ZQ, Wong SC (2004) Performance evaluation of signalized urban intersections under mixed traffic conditions by gray system theory. J Transp Eng 130(1):113–121
12. Li X, Shu W, Li M, Huang HY, Luo PE, Wu MY (2009) Performance evaluation of vehicle-based mobile sensor networks for traffic monitoring. IEEE Trans Veh Technol 58(4):1647–1653
13. Liu X, Sun D, Chang Y, Peng Z (2010, August) Traffic status evaluation based on fuzzy clustering and RBF neural network. In: 2010 seventh international conference on fuzzy systems and knowledge discovery, vol. 3. IEEE, New York, pp 1405–1408

14. Pongpaibool P, Tangamchit P, Noodwong K (2007, October) Evaluation of road traffic congestion using fuzzy techniques. In: TENCON 2007–2007 IEEE Region 10 conference. IEEE, New York, pp 1–4
15. Sommer C, German R, Dressler F (2011) Bidirectionally coupled network and road traffic simulation for improved IVC analysis. IEEE Trans Mob Comput 10(1):3–15
16. Van Lint JWC (2010) Empirical evaluation of new robust travel time estimation algorithms. Transp Res Rec 2160(1):50–59

A Cooperative Motion Control Strategy of Multi-objects Simulation Based on CAV Testing Platform

Yuning Wang, Jiahao Huang, Bo Wang, Yiming Hu, Yubin Hu and Qing Xu

Abstract Test system of Connected and Autonomous Vehicles (CAV) is a newly-developed research area. Nowadays, CAVs are basically tested in real time field test. However, the field test has two disadvantages: (i) the safety of pedestrians and other conventional vehicles cannot be guaranteed in field test, (ii) the testing period can be prolix because of the scarcity of certain extreme driving conditions. In this paper, a cooperative control algorithm of mobile unit is designed. In the proposed algorithm, motion and reaction of different traffic participants can be simulated, such as pedestrian, animals and non-motor vehicles. Motion models of traffic participants are built up in two dimensionalities: trajectory and reaction. After setting appropriate simulation parameters, the scarce driving scenarios can be easily reproduced so that the CAV test period is expected to significantly shortened. We also build up a comprehensive testing framework to realize the designed function. In the end, numerical analysis and experiments are carried out to show the algorithm's feasibility.

Keywords Test system · Multi-objects simulation · Motion control · Trajectory simulation · Reaction modeling

1 Introduction

In recent years, both automobile and artificial intelligence are among the most concerned issues. In the field of Connected and Autonomous vehicles (CAV) itself, the

Y. Wang · J. Huang · B. Wang · Y. Hu
School of Vehicle and Mobility, Tsinghua University,
Beijing 10084, China
e-mail: wangyn16@mails.tsinghua.edu.cn

Y. Hu
Department of Electronic Engineering, Tsinghua University, Beijing 10084, China

Q. Xu (✉)
State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing 100084, China
e-mail: qingxu@tsinghua.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_92

whole system has been increasingly developed, with theoretical support and practical engineering applications. However, there is a field where the current research is relatively rare, which is the test system of CAV. Up to now, no unified criterion and methodology has been formed.

For vehicles, safety is always first priority. This criterion is also suitable for CAV in the future. So how can we determine that an autonomous vehicle can really deal with complex road conditions to meet people's needs? How can we ensure that a technology is safe at the automotive level? How do we ensure that these technologies could achieve their claimed level of intelligence? All these problems above require a test system to solve. Make an analogy of holding a driving license test for CAV, a third-party platform to examine the reliability of them is needed.

At present, the CAV testing mechanism can be divided into two types: fixed field test and real road test. The method of fixed field test is to separate a piece of open space and construct various road sections, such as intersections, squats, urban sections, high-speed sections, etc., and then use simple machines to simulate tested pedestrians. The machine is shown in Fig. 1. The mechanical structure is quite simple and crude, and can only roughly indicate that someone has passed through. A manikin is placed on a small tray with wheels. Then working staff drags the small flat plate through the rope to make the model man to move. In spite of its convenience, it is obviously impossible to accurately simulate the motion characteristics of people. Precise manipulation is not available. Generally speaking, the simulation method is simple, the reference values is small, and the cost is huge. It is not a suitable and reliable test system for CAV.

Another test method: real road test, is to put the measured CAV on the road of cities, and collect data through real road conditions to conduct experiments. Although the test result is more reliable, the safety of pedestrian and other conventional vehicles are threatened, and the testing period is prolix since some specific required condition do not occur very often. Therefore, there is currently no testing mechanism that both meets the needs and guarantees the safety as well.

Fig. 1 Current closed field test method for CAV

For lack of a developed test system, we hope to combine the two existing modes, gather the advantages and avoid the defects, and design a safe, stable and reliable universal test system for intelligent and connected vehicles. Hence, we use simple robots to simulate motion of transportation participants, controlling mobile units to reproduce the needed test scenarios.

At present, there are many algorithm models for simulating human motion in the robot field, such as charge wall model and A* algorithm. However, these models are not suitable for traffic behavior simulation. There are two reasons for this: (1) The fundamentals purposes of these control models is not to simulate real pedestrians, but to optimize for certain objectives. The starting point is different; (2) Existing models are rarely modeled for traffic situations. Therefore, we decided to innovate and develop a motion simulation model suitable for simulating traffic situations.

2 Analysis of the Functions of CAV Testing Platform

Traffic situation refers to the states and situations of all traffic units that are deployed or in motion and the behavior in the drivers' interest-sensitive area and it contains all of the information about the traffic entities [1]. According to the concept and modeling of driving safety field [2], the participants of traffic in the era of automatic driving is composed of vehicles(drivers), roads, pedestrians, and other traffic factors like animals and non-motor vehicles like bicycles. To achieve the testing scenario, we must reproduce these composers. Within them, roads are easy to simulate because they are static and invariable. For vehicles, we can use multiple CAVs, take one as measured object and other as situation factors, which is not difficult to handle. The key point is the simulation of pedestrians and other moving participants. Each unit of them contains full degrees of freedom, and currently there is no existing tailor-made program or machine that can imitate moving transportation participants. Therefore, we lay our study emphasis on the imitation of pedestrian, non-motor vehicles and animals. After investigation, we decided to simulate the behavior of traffic objects by two dimensions: the trajectory dimension simulation and the reaction dimension simulation.

3 Motion Model of Transportation Participators Simulation in CAV Testing Platform

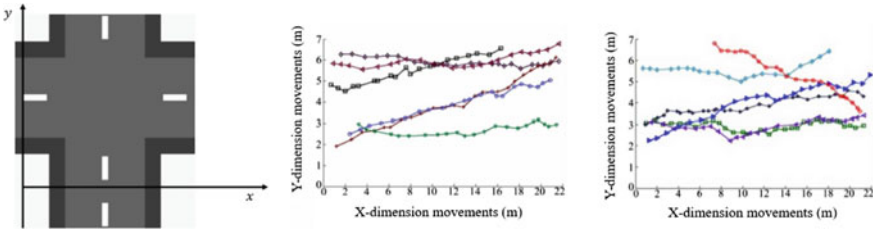
3.1 Trajectory Dimension Simulation

Motion trajectory simulation refers to emulating the speed, route, and period of a single unit. One fundamental motion feature of the living creature is that its motion is not a simple uniform one, but reciprocating with a certain period of time, the length

of which depends on the motion mechanism of the object. Taking the pedestrian as an example, the movement between his reciprocating swing arms constitutes a cycle, while with regard to the dog, the defined cycle refers to the movement between its two left fore paws touching the ground. From plenty of literature on bionics and traffic behavior, we abstract the motion patterns of a series of objects in specific traffic scenarios. Data of some species is difficult to get access to, hence we take real sampling of animals or people. Figure 2 shows movements information which was recorded in the crossing simulation.

Additionally, it is clear that the natural moving speed of each subject or species differs from each other. In addition, even the same specie will change the rate of movement in different moods, ages or physical conditions. Considering the focal discrepancies, we classify all traffic objects according to two characteristics—one is the species and the second is the mood—in order to establish a complete object simulation library, and the two form a two-dimensional object library. The difference of species is embodied by various curve shape in one cycle, and the difference of mood by the frequency of cycle and the size of the amplitude.

During the study, we divide the species into 6 groups: human, non-motor vehicles, pets, poultry, and livestock. For moods, we divide them into three categories: sluggish, average, and restive, which respectively correspond to low, medium, and high mode. Through two dimensions of population and moods, we can better match traffic objects to the corresponding characteristics. For instance, the elderly can correspond to (human, sluggish), and the young man can correspond to (person, restive); Wild cats can correspond to (cats and dogs, restive), and ordinary dogs can correspond to (cats and dogs, average). The specific object library classification can refer to Table 1 below. In each blank I raise one possible traffic participants that match the classification. For each kind of class, we unify the moving curve and codify them on base platform.



(a) Coordinate system (b) Youth (c) Mid-age

Fig. 2 Moving curve of different people when cross roads [3]

Table 1 Transportation participants library and classification

	Sluggish	Average	Restive
Human	Old Man	The mid-aged	Youth
Non-motor Vehicles	Pedicab	Bicycle	Electromobile
Pets	Kitty	Ordinary Dog	Wild Cat
Poultry	Goose	Rooster	/
Livestock	Cow	Sheep	Deer

3.2 Reaction Dimension Simulation

In terms of the dimension of consciousness, we reproduce the response when the subject receives an external stimulus. Different objects will react differently. We define these reactions as consciousness. For example, considering the scene of crossing roads, when vehicles come from the side direction while someone is at the middle of the road. Mainly there are three kinds of reaction: (1) Accelerate to pass with slightly turning aside, (2) Stop to wait until the vehicle has passed, and (3) Step back to ensure the safety. After sampling and investigation, we found that human beings are more likely to take action (2) and (3), while pets are more likely to accelerate, as shown in Fig. 3.

It should be noted that the response is not a certain event but a random and unpredictable one, so the difference in consciousness is not reflected in the response itself, but in the probability of distribution to various behaviors. Hence in order to obtain accurate behavior pattern of one subject or species, tremendous amount of data needs to be collected and analyzed in order to eliminate the randomness. With the help of the data, a behavior tendency table can be made in which every possible reaction is weighted by probability.

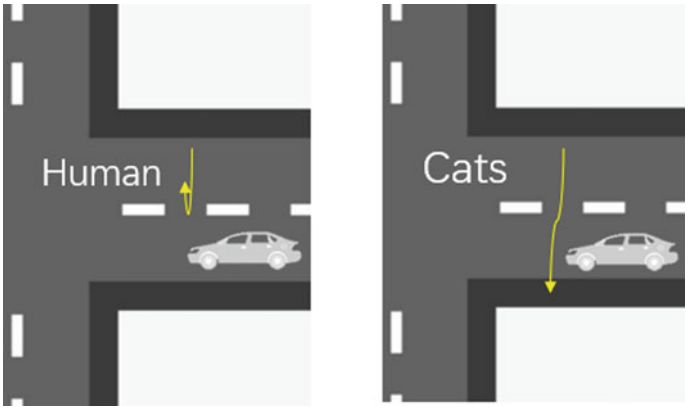


Fig. 3 Human beings' and Cats' different reaction to external stimulation

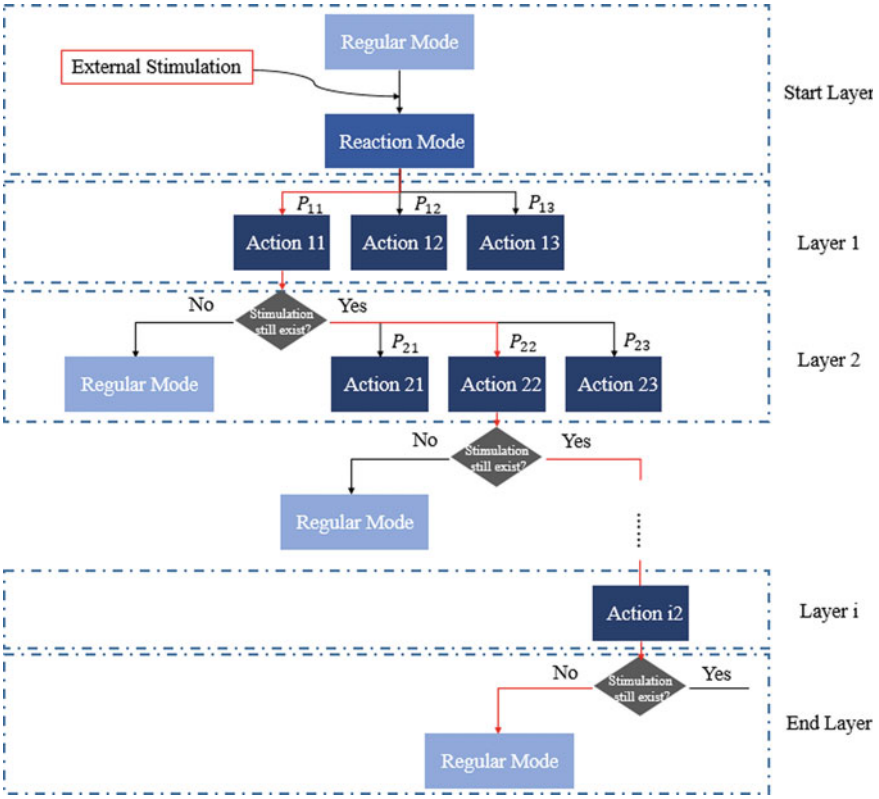


Fig. 4 Overall process of reaction simulation

Our codes also embodies the probability of this dimension of consciousness. The distribution greatly varied from specie to specie, and slightly varied from mood to mood. For a single judgement, a random number is generated to judge, and adjust the threshold of the judgment to change the different reaction modes. For the entire complete action series, we use a decision tree to implement, and the procedure will not end till the external stimulation vanishes. The overall process can be illustrated by Fig. 4.

4 Testing Framework of Platform

To fulfill fundamental functions, the system takes advantage of machines to imitate traffic participation objects, restores complex road conditions that may occur in real roads, and tests the response of intelligent vehicles in complex environments. In order to enable the motion control strategy to be applied to the CAV test platform, a support

system has been developed to implement the overall platform. The framework of the system is shown as Fig. 5.

The Appearance simulation imitates the physical shape of the traffic participants of pedestrians, animals and bicycles through mechanical structure design, and at the same time realizes posture simulation such as swinging arms through simple joints and mechanisms. At the same time, with the help of retractable joints, the size of model can be changed according to the needs. Diagrams are as shown as Fig. 6.

The communication structure, that is, the visualization of the whole system and the collection and processing of information such as speed and position, is the skeleton of the whole testing system. The overall communication network can be roughly divided into four kinds of topology structure. First, the virtual CAN bus is used to transmit information on the vehicle side, and the differential GPS data in Json format is transmitted to the vehicle terminal by ZeroMQ, and then the data is transmitted to the OBU by the udp protocol. The above is the vehicle-side communication; then the

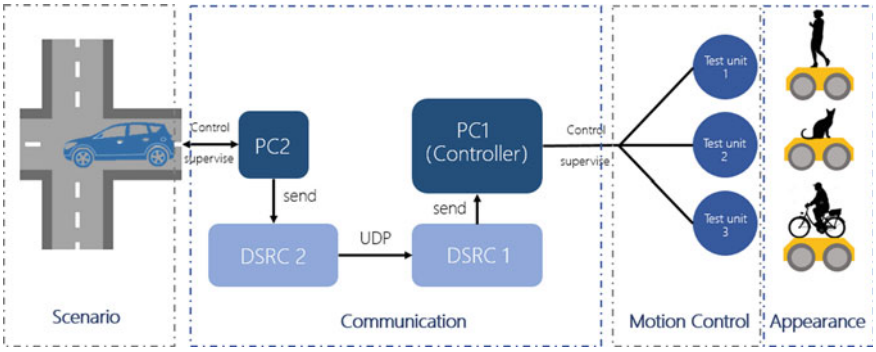


Fig. 5 Supporting system of CAV testing platform

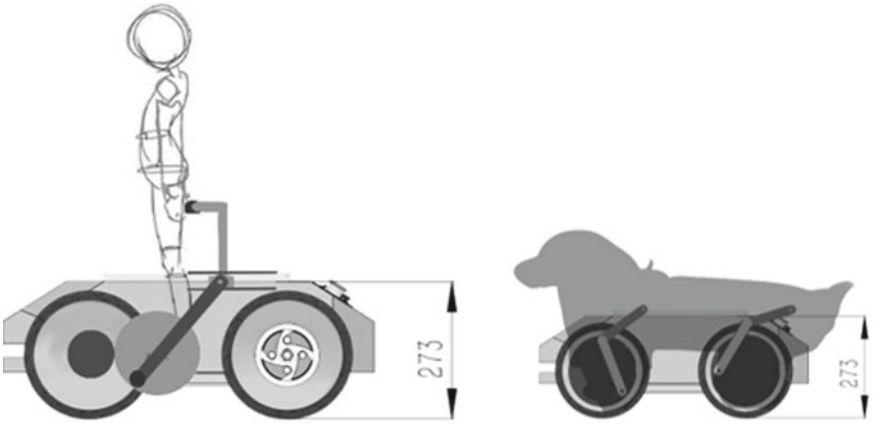


Fig. 6 Diagrams of appearance simulation

DSRC is used to realize the information transmission and reception of the vehicle under test and the controller PC. The OBU at the vehicle end forwards the previously obtained GPS information to the RSU; on the test units side, we uses the master-slave mode of Ubuntu, setting Talker and Listener ports to receive the speed of the car and control its movement; finally, all the information is gather on the controller PC, using RVIZ visualization, as shown in Fig. 7. And the Final system is presented as Fig. 8.

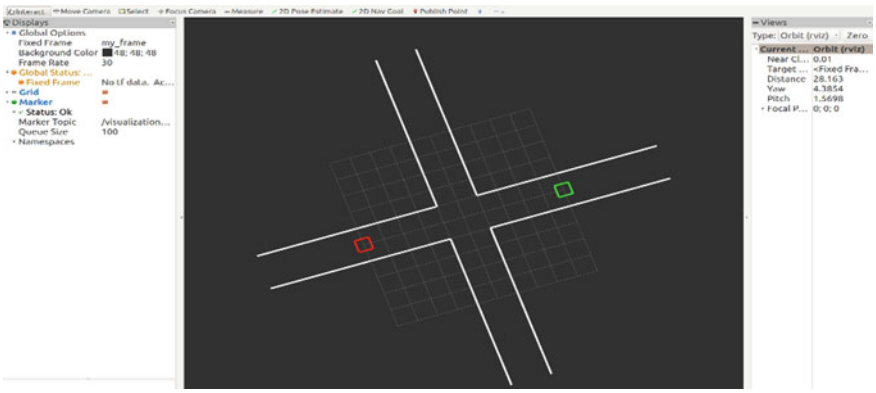


Fig. 7 Controller PC visualization

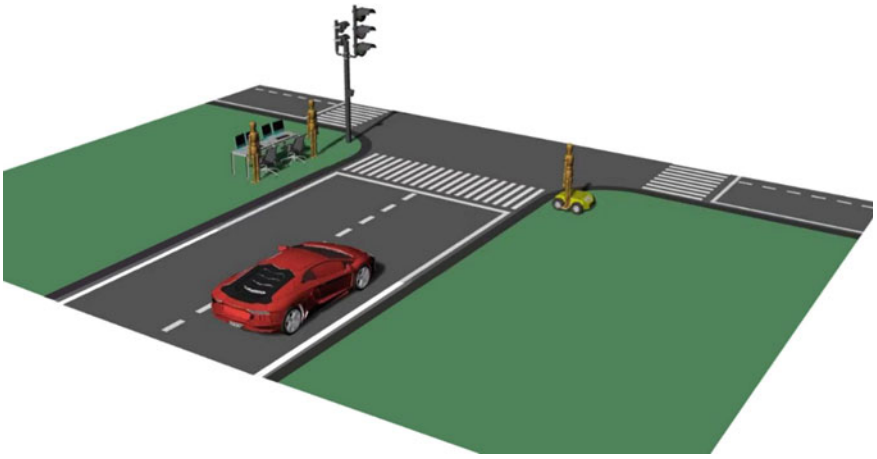


Fig. 8 Overall testing platform for CAV

5 Model Verification

In order to demonstrate the effects of our motion simulation algorithms, we conducted real-condition experiments, also taking crossing roads as an example. Firstly, a moving unit which is available for further development is needed.

The moving mechanism of vehicles and people is different. Apart from the discrepancy of velocity and physical size, the main difference is the mechanism of turning. Figure 9a, b are the schematic diagrams of vehicles' and human's turning mechanism. Due to rigid axle and the structure of power train system, when make turnings vehicles has to drive as an arc with the radius of curvature R. In contrast, human can make any broken-line turning as wish. Hence we need to select a hardware design platform which can realize this kind of move.

After searching existing robots developing platform, a moving robots named Autolabor Pro1 is employed as the base. Each wheel is connected with a motor. Turing is realized by differential steering. The mechanism is shown in Fig. 10.

Suppose that the moving velocity of robot is v_c, now the designed rotation speed is ω_c. The quantity controlled is the rotation speed of motors of both sides ω_1 and ω_2, where $\omega_i \cdot r = v_i$, r equals to the radius of wheels.

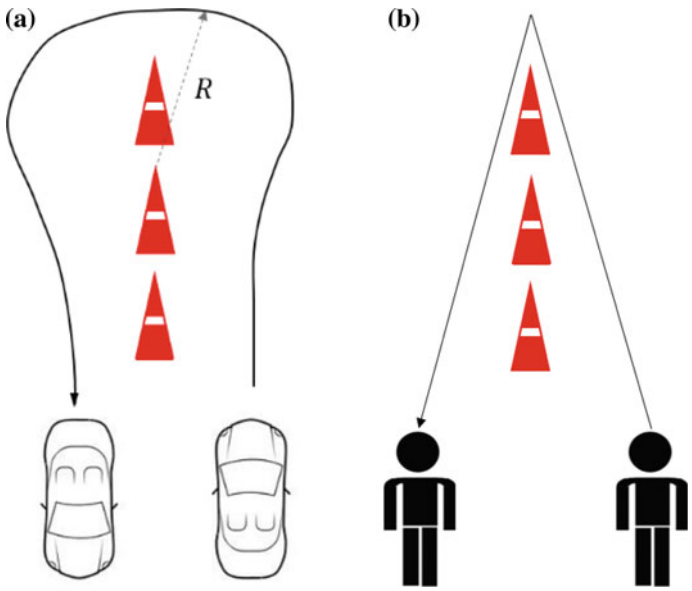


Fig. 9 Schematic diagrams of vehicles' and human's turning mechanism

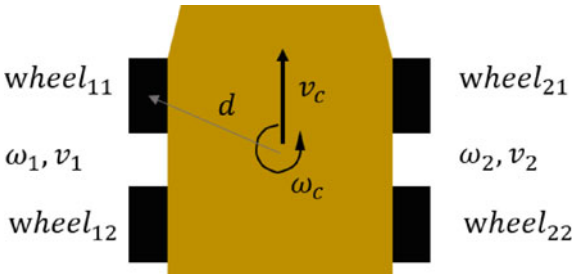


Fig. 10 Autolabor's turning mechanism

We have the follow Eq. (1)

$$\begin{cases} \omega_i \cdot r = v_i \\ v_i = v_c \pm v_{rel} \\ v_{rel} = \pm\omega_c \cdot d_i \end{cases} \tag{1}$$

Solving Eq. (1), we have the required motor rotation speed (2)

$$\begin{cases} \omega_1 = \frac{v_c - \omega_c \cdot d}{r} \\ \omega_1 = \frac{v_c + \omega_c \cdot d}{r} \end{cases} \tag{2}$$

Through controlling motor rotation speed to the answer we derive, we can realize broken-line turning and set the foundation for simulation of pedestrians and other transportation targets.

We firstly tested the feasibility of the overall system (shown in Fig. 11), then reproduced scenarios of three kinds, which are: old man, man and young man cross roads (shown in Fig. 12), corresponding to Human_Sluggish (HS), Human_Average (HA) and Human_Restive (HR) models in the libraries of codes. We also use a real Connected and Autonomous Vehicle as a measured target (shown in Fig. 13). Finally, we drew the *v-t* curve, and compare it with the trajectory of the human motion we drive before. The result is presented as Fig. 14a–f.

It can be seen that our abstracted model is reasonable. Old man can well correspond with the HS model. There are three main features of motion: the periodic variation of the peak (beat, which are shown by the curves of dark colors in figures above), the time between the adjacent peaks (period), and the peak and valley values. In a single beat, the peak change of the model and the speed is firstly smaller and then grow larger; the peak value is, on average according to sample database, 1.50 m/s for experiments and 1.55 for real pedestrians. However, the valley value is slightly different. The real minimum is 0.38 m/s and our motion model will be 0.24 m/s in each cycle, which is due to the limitations of the hardware platform, since the motor has a certain tiny delay from changing voltage to starting, and the hardware platform will be improved in the future; the duration of a single cycle is about 0.8–1.2 s, which is suitable. Therefore, it is reasonable to use the HS model to compare the behavior

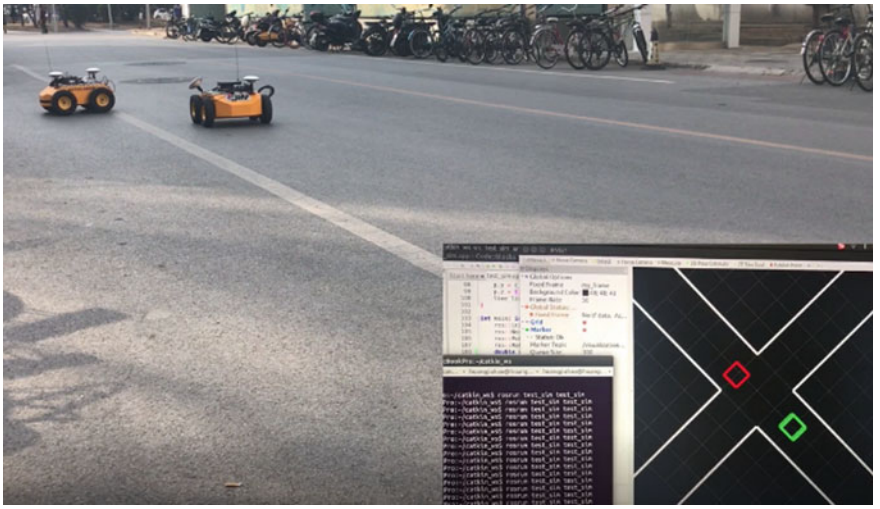


Fig. 11 System feasibility experiment

Fig. 12 Comparison experiment of motion control

of the elderly crossing the road. The result is similar for the case of man and young man.

Compared with the above six curves, although limited by hardware equipment and data volume, we can't achieve a wonderful re-engraving effect, our motion control can still roughly restore the motion law of pedestrians crossing the road.

Fig. 13 Real vehicle experiment

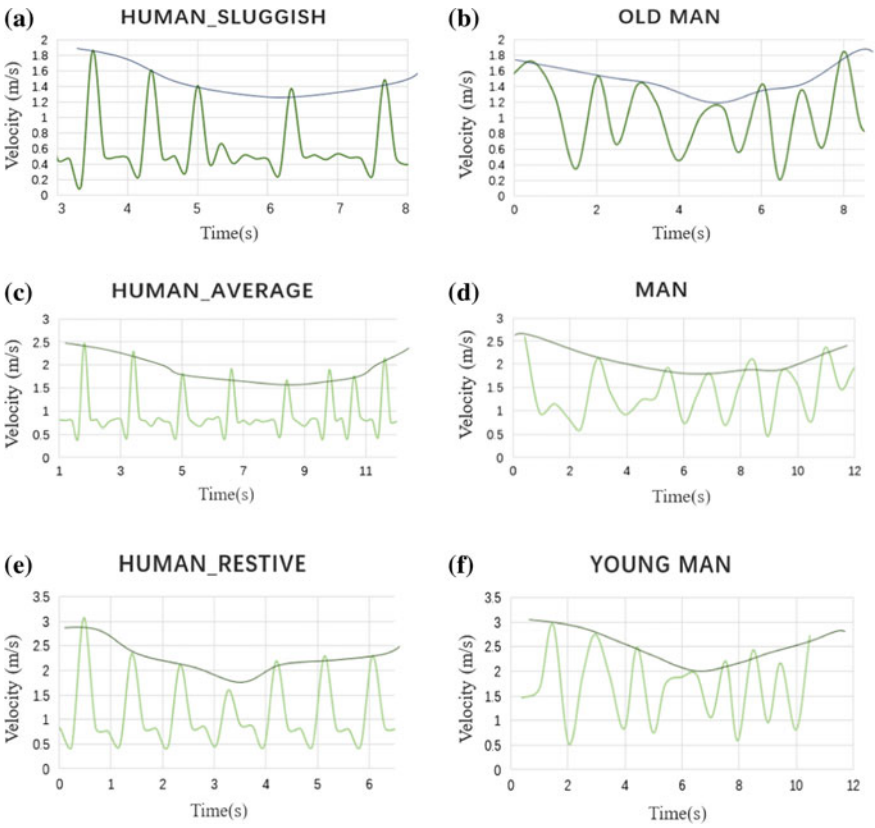


Fig. 14 **a**, **c**, **e** Results of *v*-*t* curve of motion control model experiments. **b**, **d**, **f** *v*-*t* Curve of sample data

6 Conclusions

Based on a comprehensive analysis of behavior law and the consideration of feasibility, a new cooperative motion control strategy of multi-objects simulation for CAV testing system is developed. The pivotal concept is to imitate living participants from two dimensions: trajectory dimension and reaction dimension. Trajectory points to the basic numerical value of moving features, including frequency, amplitude velocity and the shape of moving routes. We derive the origin data by both literature reading and live sampling. Reaction means the different reply to external stimulations. We use a probability distribution and decision-making tree to reproduce the whole judging process of traffic participators. In order to realize the function of this motion control strategy, we build a synthetical supporting system to make it into practical and useful. The supporting system contains a communication network, appearances imitation designs and monitoring controller that make the testing process visualized. In the end we make real-vehicles experiments to verify its feasibility and compare the actual appearance of our motion model with the original sampling data we derive, finding that our control strategy can roughly simulate the behavior of real transportation participants. Limited to hardware platform, there is still potential for progress in the future.

Acknowledgements Research supported by National Natural Science Foundation of China (Grant No. 51605245) and the Joint Laboratory for Internet of Vehicle, Ministry of Education—China Mobile Communications Corporation. We would like to extend our sincere gratitude to the supervisor of our study, Prof. Wang Jianqiang, and THICV, for all instructive advice and experiment facilities.

References

1. Lei W, Wang X, Yang X, Wang X (2007) Study on the recognition of traffic situation and its state transition mechanism. Commun Stand 2/3:61–65
2. Wang J, Wu J, Li Ya, Li K (2014) The concept and modeling of driving safety field based on driver-vehicle-road interactions. In: 2014 IEEE 17th international conference on intelligent transportation systems (ITSC)
3. Xue S (2018) Experimental and simulation study on the behavior of descendants and evacuation in several situations. Beijing Jiaotong University

Research on Attention Capacity Measurement for Drivers' Visual Space Information

Li Zhu, Jian Xiong, Fengxiang Guo and Yahui Xie

Abstract In this paper, drivers' visual space attention capacity was researched, and a way of measuring this attention capacity by driving simulation experiments has been proposed. Firstly, based on static and dynamic traffic information the quantitative and incremental stimulus information sources were established and the road traffic virtual scenes with different amounts of information sources were built. Then 30 subjects (half experienced and half novice) were selected to participate the simulation experiments, the amount of stimulus information that the subjects sensed at each test point were tested, and subjective questionnaires were carried out after each test. Finally, experiment data were statistically analyzed. The results shown that when the stimulus information was 2 or 3, the subjects could get the information, however, when the number of stimulus was 4 or 5 or 6, the amount of information that the subjects sensed were obviously decreased to 38, 19 and 8% separately. The subjects' average attention capacity was 4.27, but there was a significant difference ($p = 0.02$) between the experienced and the novice. The attention capacity for experienced drivers was 5, and that was only 3.54 for the novice. Moreover, during driving, the distribution of the amount of information obtained by the subjects was similar to the Poisson distribution. The results should have a guidance for the research of driver's psychological behavior and intervention, as well as the renovation of traffic information sources and environment.

Keywords Attention capacity · Driving simulation · Test · Traffic information · Experienced and novice drivers

L. Zhu · J. Xiong (✉) · F. Guo · Y. Xie
Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming, Yunnan 650500, China
e-mail: xjebox@163.com

L. Zhu
e-mail: 851458204@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_93

1 Introduction

According to the theory of psychological cognition, attention has capacity and can be allocated. If want to improve our attention, we must pay more special attention, there is a certain limit of human attention resources, which makes people have to control the allocation of limited attention resources to different activities or different aspects of the same activity [1]. Attention is an important psychological activity of drivers. Generally speaking, simple driving situations require less attention, while complex driving situations require more attention. For example, on a wide and void highway, experienced drivers can talk to people in the car while driving, can carry out more than one activity at the same time, because these activities do not exceed the total capacity of information processing; on the contrary, if drivers driving in a complex traffic and crowded streets can no longer distract themselves from chatting with people because of a lot of visual and auditory stimuli take up their attention.

However, the allocation of attention is not fixed, and some tasks can be automated by training, which does not require too much attention, so that more attention can be allocated to other tasks [2, 3]. Drivers have multiple tasks: observing road traffic, holding the steering wheel, controlling the throttle or braking. But through training, manipulating the car can be gradually automated, allowing drivers to focus more attention on observing the traffic. Therefore, attention is an important feature of driver's psychological behavior, which reflects the workload and attention allocation. Trained drivers realize the automation of vehicle driving, so that more attention can be allocated to observing traffic environment. The comprehensive observation of traffic environment is the basic requirement of drivers' security-conscious and an important guarantee of safe driving.

The study of drivers' attention has always been an important development direction of research on driving behavior and safety, and there are a lot of research literature, mainly focused on the driver's visual characteristics, attention allocation, attention transfer, distraction and other aspects of research [4–6], by contrast, the study of driver's attention capacity is not much. Ball et al. [7] proposed to use the UFOV (The Useful Field of View) to measure driver's attention, and obtained the relationship between attention capacity and age; later, Ball et al. [8] used UFOV to study the relationship between traffic accidents in the elderly and UFOV, indicating that the decline in attention capacity of the elderly was the direct cause of accidents. Wei and Cundao [9] tested the attention capacity of 69 bus drivers, randomly present slides with four to eleven traffic signs of random order in short time before the subjects, and records the information obtained by the subjects. According to the accident records, the subjects are divided into accident group and safety group, and the relationship between attention capacity and accident was analyzed. The results showed that the attention capacity of accident group was less than that of safety group, and the attention capacity has nothing to do with gender. Ball et al. [10] further proved the validity of the UFOV test on driver's attention capacity in 2005, and proved that the decrease of driver's UFOV would lead to the decrease of driving ability by experiments. Ariane et al. [11] studied the effect of attention capacity on driving distraction

in the elderly through a driving simulation experiment in 2015. It is indicated that the reduction of attention can alleviate the influence of distraction in driving, and the importance of the study of attention capacity in training and evaluation of driving.

The UFOV is the total visual field area in which useful information can be acquired without head and body movements; UFOV testing method is based on PC-display screen, the test signal pattern appears instantaneously on the screen, and the subjects' response complete by mouse [12]. But there is a certain difference between the attention capacity of driving and the attention capacity of traditional tests. Firstly, the information of traditional attention testing is clear, only needs to test the number of attention information the subjects got (such as the number of beans thrown out), however, the attention information of driving is related to driving or driving safety, which is selective attention information. Secondly, the appearance and disappearance of attention information of traditional testing is synchronous (generally presented by computer screen), but the process of traffic information is very short, some are synchronous, some are not; thirdly, the test information which used on screen almost does not need to search, while the information during driving requires the driver to search quickly. For drivers, the most important thing is to be able to capture as much information as possible in a dynamic environment when there are multiple signals. In this paper, driving simulation experiment is used to test the attention capacity of the subject by setting different numbers of road information at the crossroads, and try to explore the attention capacity characteristics of the drivers under the conditions closer to the real driving environment.

2 Methods

2.1 Information Sources and Testing Methods

There are a lot of information in the traffic environment, which contain static and dynamic information, these information comes from different positions, that is, different information sources. For drivers, it is important to sense information from various sources. As amount of information does not have any impact on driving, it is necessary to pay attention to the information related to route and traffic risk. Therefore, firstly, the noteworthy information should be selected from a large amount of information as the test information; secondly, the information source of each test point should be designed to form the information quantity, the time of multi-source information appearing should be shorter, and the information quantity should be increased gradually, so that the limit of information acquisition can be measured. Finally, the test points with different amounts of information are effectively arranged in the test scenario. Drivers randomly selected the order of test points in test, and whether information was identified by the subject can be collected through the questionnaire after each test point. The more information the subject noticed at the test point, the more attention the driver assigned to visual observation.

Table 1 Road information design

Information level	Test information	Test number
Level one	Traffic sign (1) + motor vehicle (1)	Two
Level two	Traffic sign (1) + pedestrian (1) + motor vehicle (1)	Three
Level three	Traffic sign (2) + non-motor vehicle (1) + motor vehicle (1)	Four
Level four	Traffic sign (2) + pedestrian (1) + non-motor vehicle (1) + motor vehicle (1)	Five
Level five	Traffic sign (3) + pedestrian (1) + non-motor vehicle (1) + motor vehicle (1)	Six

2.2 Test Information Source Settings

As the large amount of information at intersections, including motor vehicles, pedestrians, non-motor vehicles, traffic signs and so on, which are drivers need to pay attention to [13], so the experiment selects the non-signalized intersection as the experimental scene, and the motor vehicles, pedestrian, non-motor vehicles and traffic signs as the test information which drivers need to catch. In order to prevent drivers from paying too much attention to the information on the traffic sign, simple traffic signs such as "concession", "no left" and "speed limit" are selected.

The information sources are divided into five levels according to the test number, showing an increasing trend [14]. The setting of effective information is shown in Table 1.

2.3 Scene Design

The experimental scene was established by using the three-dimensional virtual modeling software VRMAG^{R}, which consists of the modeling and editing module of the virtual scene and the driving module of the visual scene. Modeling and editing module are used for the design of virtual experiment scene, and driving module is used for the display of driving simulation experiment scene.

The test scenarios of five intersections in a city are constructed according to 1–5 levels of test information, and put the five intersections into the road network to form the experimental road section (non-test intersections are randomly inserted in the road network to improve the test validity). The total length of the experimental road section is 6 km, including roads, traffic facilities and surrounding environment. The above five levels of test information were randomly set in five experimental intersections, requiring that the test information appeared and disappeared at the same time. The setting scenarios (overlook and visual spatial distribution) of various test information sources in the intersection were shown on the left side of a–e in Fig. 1. The visual spatial distribution of each information source was shown on the right side of a–e in Fig. 1.

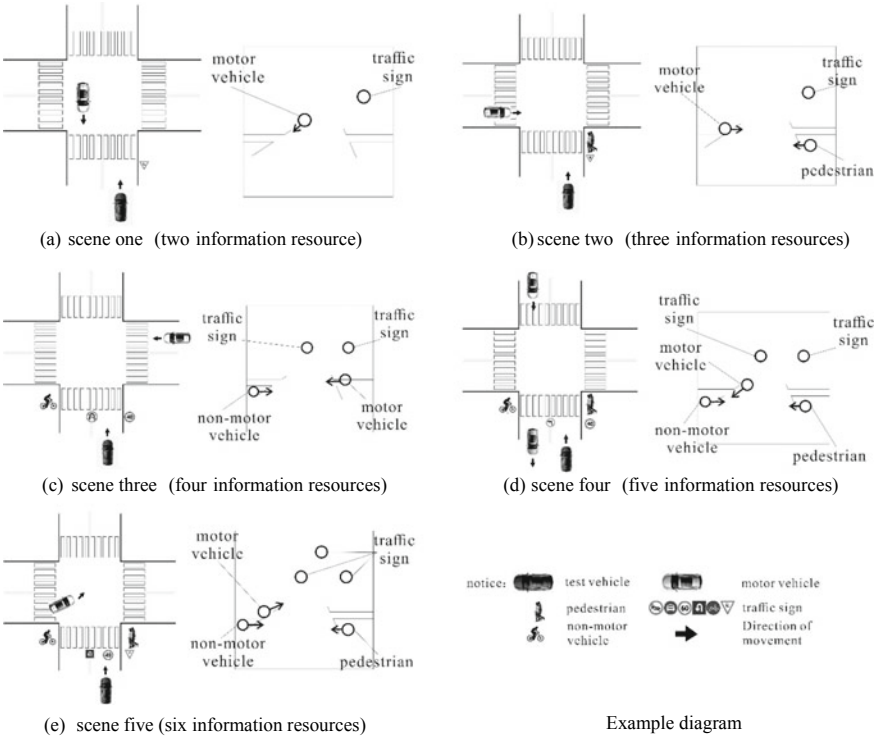


Fig. 1 Intersections with different amount of stimulus informations

The test scenarios include the above-mentioned five levels of information sources. To avoid psychological hinted to the subjects, a number of experimental scenarios were designed with different orders of the five levels of information sources, and three non-test intersections were arranged in each experimental scenario to mix with five test intersections. Each participant randomly selected one experimental scene (Fig. 2).

Fig. 2 The 5th intersection scenario

2.4 Sample

According to the driving experience, mileage, gender and other factors, 30 drivers (23–50 years old) were selected to participate in this experiment, including 18 male drivers and 12 female drivers, all with more than one year driving experience. According to driving experience and mileage, all subjects can be divided into two groups—experienced group and novice group, each group of 15 drivers, experienced drivers generally had more than five years of driving experience (mainly taxi drivers or driver's school coaches), novice drivers generally had less than three years of driving experience (mainly college students). Two subjects in each group were unable to complete the experiment because of carsickness, and the final number of effective sample was 26 (13 novice drivers and 13 experienced drivers).

2.5 Measurements

The simulation test was carried out in driving simulator KMRTS in Kunming University of Science and Technology. KMRTS is a fixed driving simulation system, which is composed of six parts: real vehicle, vehicle dynamics model, scene image generation system, traffic model system, audio system and computer control system. It can effectively complete the modeling of experiment and driving experiment [15]. The simulator has been validated in the aspects of driver's visual field, image resolution, visual range of traffic sign, speed perception, etc. [16], which can ensure the validity of the test in an acceptable range. In the experiment, the subjects randomly selected a group of experimental scenes from several experimental scenarios with different order of information sources, and carried out the simulation driving experiment on the driving simulator. The speed limit sign was set in the scene, which required the driver to drive in a specified way, keep the speed according to the speed limit requirements, and paid attention to the relevant traffic information. Each participant was required to be familiar with the driving simulator and began the formal experiment after the preliminary experiment. Record the speed, acceleration, deceleration and other driving behavior data when each participant drive through a test point. Then stopped for one minute, the experimenter was provided information survey questionnaire and attention load perception subjective questionnaire. The participant chose the noted information and filled in the attention load feeling questionnaire, and then continued to drive to the next test point until all tests were completed (Fig. 3).

The information quantity of attention referred to the amount of information that different subjects had noticed in a certain period of time. As the test was carried out during driving, different subjects' speeds were different, it would affect the time of the subjects' attention. In order to eliminate the impact of speed, definite

$$AM = \frac{SM}{t} \tag{1}$$

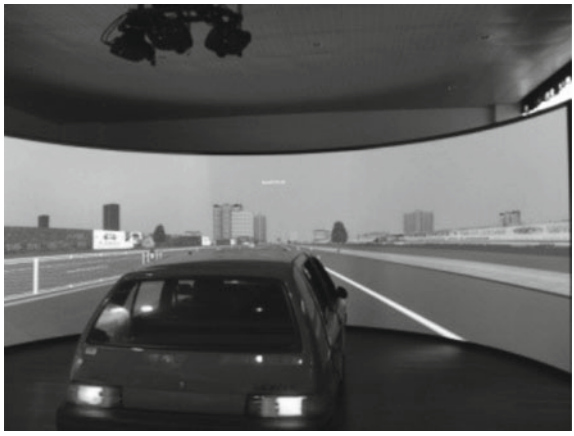


Fig. 3 KMRTDS driving simulator

In the formula, AM for attention, SM for the amount of information that the driver should pay attention to when passing the section L, t for the time that the driver pay attention to the information.

As the speed of the subjects passing through this section was V_i, the subjects' attention time was

$$t_i = \frac{L_i}{V_i} \tag{2}$$

In the form, L_i was the distance from the point that driver should notice the information to the point that driver pass the intersection. Thus, the amount of subjects' attention at the test point can be obtained:

$$AM_i = \frac{SM_i}{L_i} V_i \tag{3}$$

Through the statistical analysis of the experiment data, the amount of attention each subject got within unit time at each test point was calculated. The maximum amount of attention that the subject can obtain is the attention capacity of the subject. AC was used for the attentional capacity of subjects.

$$AC = AM_{\max} \tag{4}$$

If the experimental design was reasonable and the experimental process was effective, the amount of attention that the subject got at each test point should increase with the increasing of information provided of the test point to the capacity limit.

3 Results

The subjects were asked to keep the speed at 60 km/h. If there was a big difference between the required speed and the speed of approaching the test point, the subjects would be reminded to adjust the speed. The speed in the experiment was kept between 55 and 65 km/h. As the distance of appear and disappear of the test information in the five scenarios were the same, the subjects' cognitive time were 5–8 s at each intersection, which basically met the test requirements. The difference of cognitive time between the subjects could be neglected.

The amount of information obtained by the subjects was equivalent to a score (each person notices that one information was one point, two information was two points, and so on, up to six points). Table 2 was the total statistics. Tables 3 and 4 were the statistics of experienced group and novice group respectively. The data in the table was the number of subjects who got the score, and the last column was the score. For example, in Table 2, there were 4 subjects who noticed one test information in scene 1 and 22 subjects who noticed two test information, the total score was $1 \times 4 + 2 \times 22 = 48$ points. There were 23 subjects who noticed three information in scene 2 and 3 subjects who noticed two information, the total score was $2 \times 3 + 3 \times 23 = 75$ points, and so on.

The difference of the amount of information obtained by the two groups in the same scenario was analyzed. The results were shown in Table 5. Taking the number of test information as abscissa, the average number of the information the subjects obtained as ordinate, and the variation trend of number of the information different subjects obtained (total sample, experienced group, novice group) was shown in Fig. 4. It can be seen that when the amount of test information was relatively small, the subjects could all notice the information. As the amount of test information increased, the amount of information the subjects obtained began to decrease, which was consistent with the cognitive characteristics. There was a significant difference between the experienced group and the novice group in the scene 3–5 (Table 5): the amount of the information the experienced group obtained was significantly higher than that of the novice group. According to Eq. 4, attention capacity was the maximum attention, so the average attention capacity of experienced group and novice group was 5 and 3.54 respectively from Tables 3 and 4. From Table 2, the average attention capacity of total sample was 4.27.

Quadratic function was used to fit the relationship between the number of the information the total sample obtained and the number of test information. As shown in Fig. 5, the fitting function was as follows:

Table 2 Number of subject who obtain the amount of information in the scene for all samples

Test scene	Amount of information obtained						Total subject	Total score	Average score
	1	2	3	4	5	6			
Scene 1	4	22					26	48	1.85
Scene 2		3	23				26	75	2.88
Scene 3		6	10	10			26	82	3.15
Scene 4		1	1	19	5		26	106	4.07
Scene 5	1		4	9	10	2	26	111	4.27
Number of subject scored	5	32	38	38	15	2	$\Sigma = 130$		
Total score	5	64	114	152	75	12		$\Sigma = 422$	3.25

Table 3 Number of subject who obtain the amount of information in the scene for experienced group

test scene	Amount of information obtained						Total subject	Total score	Average score
	1	2	3	4	5	6			
Scene 1		13					13	26	2
Scene 2			13				13	39	3
Scene 3			5	8			13	47	3.62
Scene 4				8	5		13	57	4.38
Scene 5				2	9	2	13	65	5
Number of subject scored	0	13	18	18	14	2	$\Sigma = 65$		
Total score	0	26	54	72	70	12		$\Sigma = 234$	3.60

Table 4 Number of subject who obtain the amount of information in the scene for novice group

Test scene	Amount of information obtained						Total subject	Total score	Average score
	1	2	3	4	5	6			
Scene 1	4	9					13	22	1.69
Scene 2		3	10				13	36	2.77
Scene 3		6	5	2			13	35	2.69
Scene 4		1	1	11			13	49	3.77
Scene 5	1		4	7	1	0	13	46	3.54
Number of subject scored	5	19	20	20	1	5	$\Sigma = 65$		
Total score	5	38	60	80	5	0		$\Sigma = 188$	2.89

Table 5 Significant difference analysis of attention

	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5
Average score of the experienced group	2	3	3.62	4.38	5
Average score of the novice group	1.69	2.77	2.69	3.77	3.54
Significance	0.030	0.070	0.001	0.009	0.000

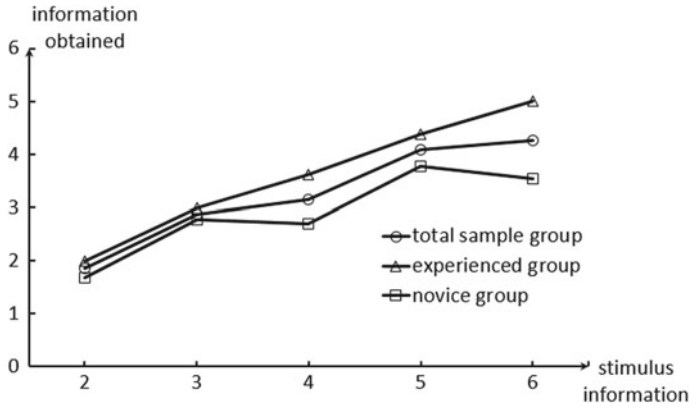
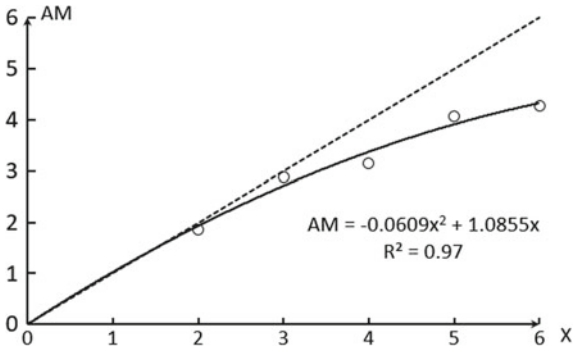


Fig. 4 Relationship between information obtained and stimulus information

Fig. 5 Function of the attention and information



$$\mathrm{AM} = -0.0609\mathrm{x}^2 + 1.0855\mathrm{x} \tag{5}$$

$$\mathrm{R}^2 = 0.97$$

The dotted line in Fig. 5 shows that the amount of information obtained is equal to the information quantity of attention. It could be seen that when the average number

of information exceeded 3, the attention curve began to move away from the dashed line.

From Tables 2, 3 and 4, we could see that the total times of subject tested in this experiment was 130, including 65 for experienced group and 65 for novice group; the total number of subject obtaining two information in all five scenarios was 32, including 13 for experienced group and 19 for novice group; the total number of subject obtaining three information in all five scenarios was 38, including 18 for experienced group and 20 for novice group, and so on. we could calculate the proportion (frequency) of the experienced group who obtain two information in the total number of experienced group $P = 13/65 = 0.2$, the novice group was $P = 19/65 = 0.29$, the total sample was $P = 32/130 = 0.24$, so as to calculate the frequency of the subjects obtain different amounts of information in the experiment, as shown in Table 6, the frequency distribution of the amount of information obtained by the subjects was as shown in Fig. 6.

From the distribution form, the frequency distribution of the subjects' attention was similar to Poisson distribution, and the average attention of the three groups were the total sample group $\lambda = 3.25$: experienced group $\lambda = 3.60$: novice group $\lambda = 2.89$. The distribution density function of the above three groups could be obtained as follows:

$$p(x) = \frac{e^{-3.25}3.25^x}{x!}, \quad x \in \{1, 2, 3, 4, 5, 6\} \tag{6}$$

Table 6 Frequency at which the subjects obtain information

	The amount of information the subjects obtained					
	1	2	3	4	5	6
Total sample	0.04	0.25	0.29	0.29	0.12	0.02
Experienced group	0	0.2	0.28	0.28	0.22	0.03
Novice group	0.08	0.29	0.31	0.31	0.02	0

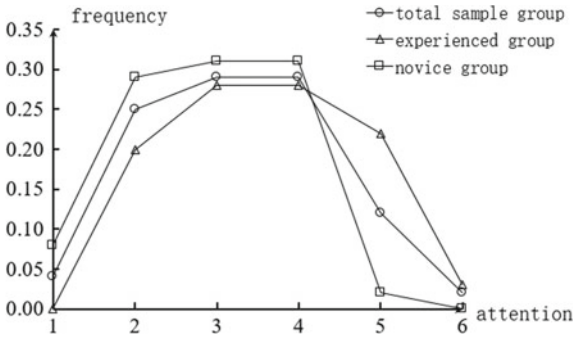


Fig. 6 Frequency distribution of subjects' attentions

$$p(x) = \frac{e^{-3.60}3.60^x}{x!}, \quad x \in \{1, 2, 3, 4, 5, 6\} \tag{7}$$

$$p(x) = \frac{e^{-2.89}2.89^x}{x!}, \quad x \in \{1, 2, 3, 4, 5, 6\} \tag{8}$$

The distribution density of drivers' attentions was shown in Fig. 7. Distribution fitting test and confidence level was $\alpha = 0.05$.

As $\lambda = 3.25$: $X^2 = 18.171 > X_{0.05}{}^2(5) = 11.070$;

As $\lambda = 3.60$: $X^2 = 9.551 > X_{0.05}{}^2(4) = 9.488$;

As $\lambda = 2.89$: $X^2 = 16.197 > X_{0.05}{}^2(4) = 9.488$;

The subject need to run their eyes over the experiment scene when they pay attention to traffic information, so there was a certain order of information acquisition. According to the questionnaire statistics after the experiment, the order of attention between the experiential group and the novice group was as shown in Table 7.

As can be seen from Table 7, in the process of the attention of traffic information, experienced group first observed pedestrians or non-motor vehicles, while

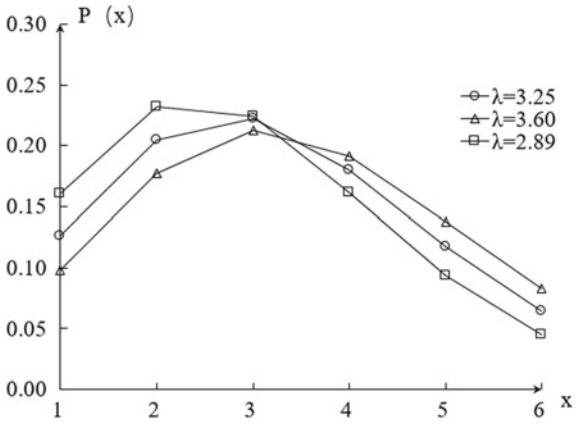


Fig. 7 Distribution density of drivers' attentions

Table 7 Attention order of subjects' information

	Order/proportion					
	First attention	Proportion (%)	Second attention	Proportion (%)	Third attention	Proportion (%)
Experienced group	Pedestrian or non-motor vehicle	69	Motor vehicle	62	Traffic sign	62
Novice group	Motor vehicle	62	pedestrian	69	Traffic sign	69

novice group first observed motor vehicles, and there were certain differences in risk awareness between the two groups.

After each test was completed, the Likert Psychological Test Scale was used to conduct the subjective observation of attention load test, and the scale was divided into five levels from weak to strong. For Scenario 1 to Scenario 3, there was no significant difference between the experienced group and the novice group, but for the more complicated Scenario 4 and Scenario 5, the average load perception of the experienced group was 3.08 and 3.31, the average load perception of the novice group was 3.54 and 3.92, respectively. There was significant difference between the two groups ($p = 0.02$), that was, the average attention load of the novice group to Scenario 4 and Scenario 5 sensitivity was higher than that of the experienced group.

4 Discussion

The results showed that, the percentage of information quantity obtained in total information was reduced with the increase of the amount of information when driving, and the maximum attention of different drivers in a short time was different. In this experiment, the average attention of subjects was 4.27. In the early soybean sprinkling experiment, the attention was between 5 and 9, while the attention of the traffic signs of bus drivers was between 4 and 5. The three test results were different, indicating that attention was different while different information or different environment. In the experiment of sprinkling beans, the attention information of the test was the same object (beans); in the traffic sign test, the attention information of the test had different shapes and colors, and the information was more difficult, so the attention was reduced; in this experiment, the attention information of the test included pedestrians, non-motor vehicles,motor vehicles and traffic signs. What is more important was that the subjects can not only specialize in attention test objects, that is, the subject had part of the attention to drive the car. But the test information sources were distributed in a larger visual space, the driver must scan to pay attention to the relevant information in a short time, so the amount of attention would drop. Therefore, the attention tested in this experiment can also be called the attention of driving. It is easy to explain why experiential subjects had more attention than novices (the average was 5 in the experiential group and 3.54 in the novice group), because experiential subjects tend to be more automatic about vehicle control, they had less attention allocated to driving than novices, so they can pay more attention on the search of information in traffic environment, more information can be noticed. The subjective attention load questionnaire also showed that the attention load of novice group was greater than that of experiential group when the amount of information increased (such as scenario 4 and scenario 5), indicating that the visual attention of novice group was less than that of experiential group.

From Tables 2, 3 and 4, we can see the relationship between driver's attention and the amount of information. When there were only two or three test information, the subjects could almost notice, of which experienced group were 100% and novice

group were 86%; when there were four test information, the proportion of could notice all the information was 38%, of which experienced group were 62% and novice group were 15%; when there were five test information, the proportion of could notice all the information was 19%, the experiential group were 38% and the novice group were 0%; when there were six test information, the proportion of could notice all the information was 8%, the experiential group were 15% and the novice group were 0%. Another noticeable phenomenon was that, with the increase of the amount of information, although the proportion of subjects noticed all the information decreased significantly, the total amount of noticed information increased, and the distribution of the amount of obtained information by the subjects in the experimental scene was similar to the Poisson distribution. In the experiment, the average amount of information the novice group can noticed was less than 3 ($\lambda = 2.89$); the amount of information the experienced drivers noticed was close to 4 ($\lambda = 3.60$).

From the order of attention, the experiential group first paid attention on pedestrians or non-motor vehicles, followed by motor vehicles and traffic signs; the novice group first focused on motor vehicles, followed by pedestrians, non-motor vehicles and traffic signs. In the traffic environment, both pedestrians and non-motor vehicles belong to vulnerable group. Once the traffic accident happened, the injuries of vulnerable group were greater, so the pedestrians and non-motor vehicles needed more attention on. Therefore, experienced drivers would first pay attention to pedestrians or non-motor vehicles, followed by motor vehicles. On the other hand, compared with motor vehicles, pedestrians or non-motor vehicles were not easy to be noticed because of their small size, so they should be an important target of visual search. The different choice of the attention of information between novice group and experienced group reflected their lack of experience and safety awareness.

The test of attention based on driving simulation is closer to the real environment. The variety of test information, the attention allocation and attention sequence of the subjects can be reflected in the experimental conditions and results. Therefore, compared with the traditional attention test method, the attention test based on driving simulation can be better on show attention characteristics of the driver and its impact on traffic safety. However, there are some limitations and shortcomings in this experiment. Although the questionnaire was filled immediately after each test point tested, but affected by short-term memory, the subjects will have certain deviations in recalling information. The problems need to be further improved in future research.

5 Conclusion

(1) The method of attention measurement based on driving simulation is closer to the real situation. The diversity of information sources, the characteristics of attention distribution and attention sequence of subjects can be reflected in the process of the test. The results can better show the behavior characteristics of drivers and the impact on driving safety.

(2) When the test information was 2 or 3, the majority of drivers can notice them; when the test information was more than 3, the amount of information that the driver noticed was significantly reduced. When the number of stimulus was 4 or 5 or 6, the amount of information that the subjects sensed were obviously decreased to 38, 19 and 8% separately. The relationship between attention and the amount of information conformed to the law of quadratic curve.
(3) The average attention of the subjects was 4.27, and there was a difference between experienced group and novice group in attention. The attention of experienced drivers was 5, while that of novice drivers was 3.54. There was a significant difference between the two groups, indicating that experienced drivers could allocate more attention on observing traffic information on roads.
(4) With the increase amount of information, although the ratio of the information the driver obtained to all information decreases, the total amount of information obtained by the driver increases, and closer to the Poisson distribution.
(5) During driving, experienced drivers would first pay attention to pedestrians or non-motor vehicles, while novice drivers would first pay attention to motor vehicles, reflecting the difference in safety awareness between the two groups.
(6) Compared with the traditional attention measurement method, the attention measurement based on driving simulation in this paper is closer to the real environment. The results have guiding significance for the study of drivers' psychological behavior and intervention, as well as the regulation of the information in traffic environment.

References

1. Xiaowen L, Ling Z, Rongsheng T (2003) Modern psychology, 1st edn. East China Normal University Press
2. Danling P, Biyin Z (2004) Cognitive psychology, 1st edn. Zhejiang Education Publishing House
3. Huang L, Pashler H (2005) Attention capacity and task difficulty in visual search. Cognition 94:B101–B111
4. Geoffrey U et al (2002) Visual search while driving: skill and awareness during in section of the scene. Transp Res Part F 5(2):157–167
5. Ma Y, Wang Y (2009) Attention assignment characteristic of driver and its influences on driving safety. J Traffic Transp Eng 9(4):114–117
6. Salaheddine B, Khalid A (2010) The role of roadside advertising signs in distracting drivers. Int J Ind Ergon 40:233–236
7. Ball KK et al (1988) Age and visual search: expanding the useful field of view. J Opt Soc Am A 5(12):2210–2219
8. Ball KK et al, Visual attention problems as a predictor of vehicle crashes in older drivers. Invest Ophthalmol Vis Sci 34(11):3110–3123
9. Wei S, Chundao H (1994) Research on drivers' attention span. J Traffic Chin J Appl Psychol 4:12–17
10. Ball KK et al (2005) Predicting driving performance in older adults with the useful field of view test: a meta-analysis. In: Proceedings of the third international driving symposium on human factors in driver assessment, training and vehicle design, pp 51–57

11. Ariane C et al (2015) Does attention capacity moderate the effect of driver distraction in older drivers? Accid Anal Prev 77:12–20
12. Mouhairong (2010) Visual attention during driving: implement of the test of useful view of field. Southwest University
13. Julia W, Mark V (2012) What does the driver Lookat? The influence of intersection characteristics on attention allocation and driving behavior, accident analysis and prevention. Accid Anal Prev 45:610–619
14. Yahui X (2015) Research on drivers' attention capacity and the correlative factor. Kunming University of Science and Technology
15. Jian X, Jiguo Z, Li D et al (2002) Application and research of vehicle driving simulator for road traffic problems. China J Highway Transp 15(2):117–121
16. Liang C, Jian X, Fengxiang G et al (2017) Image calibration method of driving simulator based on vision and cognition. China J Highway Transp 30(1):129–135

An Improved Convolutional Neural Network for Monocular Depth Estimation

Jing Kang, Anrong Dang, Bailing Zhang, Yongming Wang, Hang Su, Fei Su, Tianyu Ci and Fangping Wang

Abstract Depth estimation from monocular image plays an essential role in artificial intelligence, which is one of the important ways for sensing the operating environment in automatic-driving system or advanced driving assistant system. The most recent approaches have gained significant improvement for depth prediction based on convolutional neural networks (CNNs). In this paper, a novel framework of CNNs is proposed for monocular depth estimation based on deep ordinal regression network (DORN) and a U-net structure. The new model is trained, verified in process and tested on 5000 images from a simulation experiment platform provide by "Grand Theft Auto". To eliminate or at least largely reduce the impact from ground truth with no depth values, three different training strategies were employed for network optimization. We developed an effective weighted training strategy for depth prediction to improve the estimation accuracy. The comparison of evaluations over our results and DORN demonstrated the effectiveness of our method. The results showed that the proposed method achieved state-of-the-art performances.

Keywords Monocular depth estimation · CNN · Ordinal regression · Automatic-driving system · Image classification

J. Kang · Y. Wang · H. Su · F. Su · F. Wang
China Transport Telecommunications & Information Center, Beijing, China

J. Kang · A. Dang (✉)
Department of Urban Planning, School of Architecture, Tsinghua University, Beijing, China
e-mail: danganrong@126.com

Institute for China Sustainable Urbanization, Tsinghua University, Beijing, China

B. Zhang
Department of Mechanical & Electronic Engineering, China University of Mining & Technology, Beijing, China

T. Ci
College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_94

1 Introduction

Depth estimation from single monocular has been a significant way in understanding 3D scene geometry, which has great application potential in the field of Advanced Driving Assistant System (ADAS) and automatic driving in the Intelligent Transportation Systems (ITS). Depth estimation from 2D image has been an ill-posed and challenging problem for many years. Traditional methods mostly rely on hand-crafted features which has strong interpretability in the physical sense, such as texture and perspective information, object sizes and locations [1–4]; however, the requirements for various external conditions of these methods lead to limitations and restrictions for many specific applications so that their performances have not been very satisfying. As one important branch of Artificial Intelligence (AI), computer vision has made breakthroughs due to the fast development of Deep Learning (DL) methods in recent years, from which it could be found that the Convolutional Neural Networks (CNNs) have achieved state-of-the-art performance in image segmentation, object recognition and detection, etc. Therefore, more and more studies have applied CNNs to the monocular image depth estimation problem.

Although researchers have started to explore machine learning methods to estimate depth from a single image for over ten years, most recently it has reached a new high-level progress since 2014, due to the fast development of GPU computing power and the explosion of massive image/big data. Eigen [5] propose a multi-scale deep neural network to solve the problem of depth prediction in 2014. It was expected that the connection to the Fine network after the pre-trained Coarse network can make the output of the neural network get better results; then this method had been improved by increasing the number of multi-scales, and using pooling layer and up-sampling to replace the previous Fine network in the structure. Liana [1] used a fully convolutional residual networks (FCRN [2]) upon ResNet-50 to estimate the monocular image depth, in which the fully-connected layer was replaced by the up-sampling blocks. It requires rather fewer parameters and training data using FCRN with better performance on overall depth estimation, but with more detail's loss comparing to Eigen's method [3]. Wang [4] proposed a unified framework based on the consistency between depth and semantic prediction, which incorporate both global and regional CNN respectively. Li [5], Cao [6] and Li [7] have formulated depth estimation as a multi-category classification task based on deep residual networks. Considering the continuous characteristic of the depth values, Liu [8], Li [9] and Xu [10] explored the effectiveness of the method combing deep CNN with continuous conditional random field (CRF) for depth estimation. Fu developed a deep ordinal regression network by training the network using an ordinary regression loss, which makes the effectiveness of MDE getting a new level [11]. Besides, Chen built a new dataset consisting of images in the wild with relative depth annotations, and proposed a single deep network algorithm that learns to estimate metric depth using annotations of relative depth [12].

All the mentioned above CNN-based methods usually require a large number of input pictures as the training constraint, which is also often referred to as "supervised

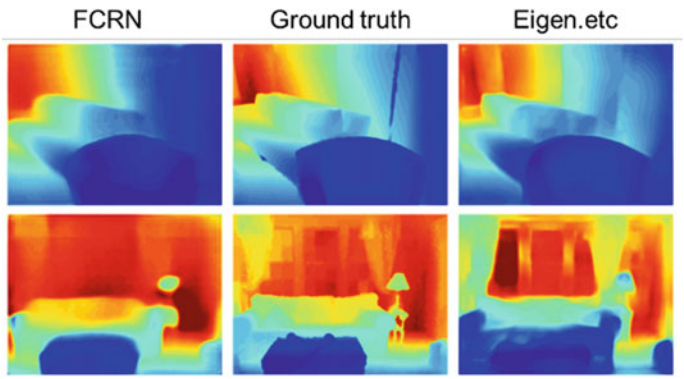


Fig. 1 Depth prediction on NYU depth results from Laina [1]

learning". Relatively, the so-called "unsupervised learning" has been explored to use the input picture training without knowing the ground truth to get the depth information, which take advantage of some physical laws between depth information and the scene [13–15]. However, the unsupervised methods with lack of reference ground truth values are often less effective than that of supervised classification currently.

In summary, the most recently development of using deep learning to estimate depth, can be divided into several categories: (1) CNN-based network: multi-scale CNNs, deep CNNs and Fully Convolutional Networks(FCN); (2) CNN + image feature: Semantic segmentation and image classification; (3) CNN + CRF; (4) Relative depth; (5) Unsupervised CNN; (6) Others, improvements of CNN architecture based on existing different networks and models. However, it could be found that these methods have unbalanced performances on the overall depth prediction and detail retention, as Fig. 1 shown. How to get better overall results and keeping the details as much as possible? For this goal, we combined two existing models to propose a new CNN architecture, using the simulation platform experiment to demonstrate the effectiveness of our method.

2 Data and Method

2.1 Training and Test Dataset

Most of the existing datasets for Monocular Depth Estimation (MDE) training and test are collected from sparse laser scan data, (i.e. outdoor benchmarking datasets, Make3D, NYU V2 and KITTI.), which makes CNNs operating on the location of missing data with corresponding poorly performance. The KITTI dataset, which was developed by Karlsruhe University of Technology, contains a total of 56 scenes

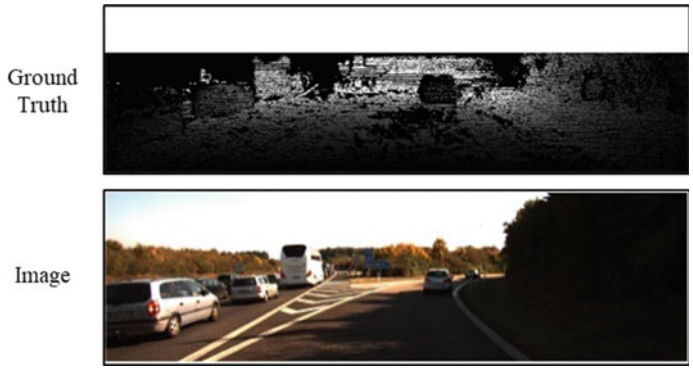


Fig. 2 Example of KITTI dataset. Depth values in the white part (sky) of ground truth are not provided by KITTI

acquired by vehicle-equipped stereo cameras. As KITTI collects traffic scenes, it has been used by a large number of research work related to transportation, including monocular image depth estimation, semantic segmentation, and vehicle detection and tracking. However, it should be noted that only the ground half of each image has a sparse depth value scanned by the lidar. The usage of pixels with no depth values usually could be masked, or the default pixel of the depth value could be filled using some interpolation or filling algorithm. Although it is still the most popular datasets for MDE, this sparse input has brought extra calculations and affected the accuracy of the results [1] (Fig. 2).

As the development of Artificial intelligence, researchers have used simple games as the simulation platform for testing algorithms. In an artificial simulation environment, it can be more easily to collect large-scale annotation data, comparing to collect and label images in the real world. Here, the evaluation of our method was completed through simulation experiments. The computer game "Grand Theft Auto" (GTA) has provided a simulation experiment platform, creating a realistic and various road environment with different vehicles driving on the road. GTA has become popular used for Autonomous Driving test, employed by the Intel Labs, the University of Darmstadt in Germany, and the "DeepDrive" project in Princeton University [16].

2.2 U-NET Based DORN for Monocular Depth Estimation

The general CNNs used for MDE mainly include convolutional layer, active layer, pooling layer, fully connected layer, loss layer, etc. The ordered superposition of these operating layers constitutes an end-to-end deep learning model. As CNNs have become increasingly popular in MDE, several improvements and new network structures have been developed. Each structure has advantages and limitations. In

this paper, an improved framework is proposed based on U-shaped [17] architecture combined with DORN [11].

The U-net architecture has major computational advantages in that it usually requires very few annotated images and uses a reasonable training time. It extended the fully convolutional network that designed the "up-sampling part", yielding a U-shaped architecture.

DORN, is a new successful DCNN-based with ordinal regression and classification thoughts, which was developed for MDE in 2018. It consists of a dense feature extractor, multi-scale feature learner, cross channel information learner, a full-image encoder and an ordinal regression optimizer. Especially, it formulated MDE as an Ordinal Regression (OR) problem, and employed the ordered information between the discrete labels that differed from typical multi-class classification method. OR is a statistical analysis model used to predict the ordered tag variables corresponding to the target. It is found that depth values have a strong ordinal correlation since they form a well-ordered set [11]. Based on the general classification model, OR model requires the serial number of labeled images. And it also differed from the regression model that the original continuous regression model prediction result would be turned into a discretized variable using OR model (Fig. 3).

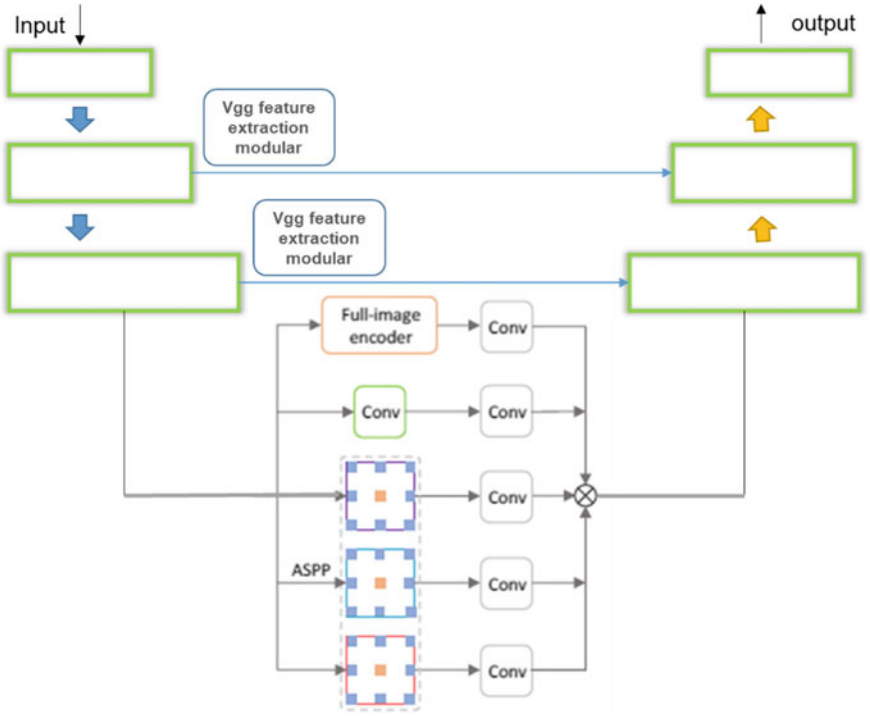


Fig. 3 Illustration of the network architecture based on DORN and U-net

3 Experiments and Results

To train a powerful CNN, it requires a certain number of samples for training and test. Here, a number of experiments were conducted for examining different aspects of our method. All the image from GTA with resolution of 1920 × 1080 captured by single monocular and depth sensors in a driving car. The scenes from "road", "city", "residential" and "Campus" categories were selected as the same with DORN as our training/test sets. The original images were down-scaled and a random crop of size 480 × 360 were used for our model training. Considering the actual needs of autonomous driving in real-world, a pre-defined center cropping following with the depth ranging from 0 to 40 m were set for evaluating our results, in which the minimum rating is 0.5 m. According to the theory of our method, the depth information of images would be divided into 80 ordered levels, and there are orderly relationships between the 80 levels. It is noted that the model was trained on the full depth range, and was tested on data with various depth ranges. Based on GTA platform, the total number of 5000 images were employed for training (3500 images, 70%), verification in process (500 images, 10%) and test (1000 images, 20%) (Fig. 4).

The algorithm is based on Keras [18] and TensorFlow [19] open source DL framework and Python 3.6 [20] for programming. All experiments and test code are run on the same computer platform. The hardware configuration includes Intel i7 3.4 GHz CPU, 16.0 GB memory, GeForce RTX 2080 Ti graphics card, and 8G video memory. The operating system is Ubuntu 18.04. The CUDA [21] version number used to accelerate the calculation is 9.0.

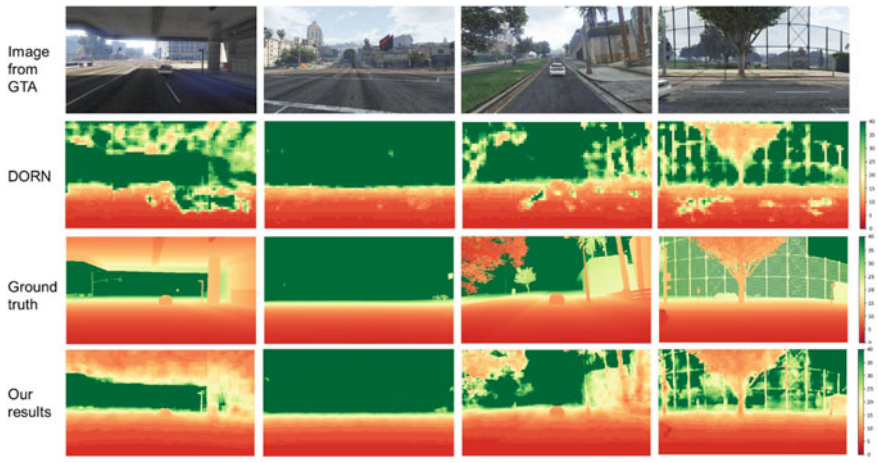


Fig. 4 Depth prediction on GTA simulation platform. The input image, DORN, ground truth and our results from improved DORN + U-NET in full-image with weight training. Pixels with distance >40 m in the image are treated as the maximum value set of 40 m

4 Discussion and Comparison

To demonstrate the effectiveness of our method, we employed the three typical indicators for assessment of MDE results: Mean absolute error (MSE), Relative absolute error (absErrorRel) and Scale invariant logarithmic error (SLOG). The following represent the formula for each evaluation indicator.

$$\mathbf{MAE} = \left|\overline{y - y'}\right| \tag{1}$$

$$\mathbf{absErrorRel} = \left(\frac{\left|\overline{y - y'}\right|}{y}\right) \times 100 \tag{2}$$

$$\mathbf{SLOG} = \mathrm{D}(y, y') = \frac{1}{n}\sum_i d_i^2 - \frac{1}{n^2}\left(\sum_i d_i\right)^2, d_i = \log y_i - \log y_i^* \tag{3}$$

where, y represent the ground truth value, and y' represent the depth prediction results.

From the existing research, the ground truth values from the existing dataset, such as KITTI as shown in Fig. 2, are not contain the values of the far distance. These pixels with no depth values usually were masked, or filled using certain algorithms, which will affect the accuracy to some extent. To demonstrate the effectiveness and accuracy of depth prediction of our method, three different training strategies were designed: Full-image training, Ground-only training, and weighted Full-image training. Full-figure training is a regular training mode, as the same as the training strategy used by DORN. Ground-only training removes the effects of far-distance, in which only half of the image (1/2) that considered as ground would be trained. For autonomous driving scenes, it is considered that depth estimation from a certain distance from the car is effective. The weighted training is based on the other two modes, to set the weight of the sky half with 0.5 for training. For the three training strategies, the corresponding indicator-G represent that only the ground half of images were assessed. As shown in Table 1, the results show that weighted Full-image training is more effective than other training methods. Although the indicators of ground-only training had the lowest values, it showed that no depth values of the far distance do

Table 1 Accuracy assessment and comparisons

GTA		MAE (m)	MAE-G (m)	SLOG	SLOG-G	absErrorRel (%)	absErrorRel-G (%)
DORN (Full-image)		5.728	3.483	11.597	11.920	31.334	33.463
DORN + UNET	Full-image	1.201	0.814	1.666	1.235	7.323	7.716
	Only-ground	–	0.596	–	0.789	–	6.118
	Full-image with weight	1.194	0.725	1.652	1.020	7.186	7.044

have a certain impact on the results of deep prediction. For the full-image training, our results are also better than DORN, which demonstrates the effectiveness of the improved network in this paper.

5 Conclusion

In this paper, we have developed an improved CNNs based on DORN and U-Net for monocular depth estimation. The new method consisting of a novel framework of CNN and some effective strategies for network optimization. The GTA simulation platform has provided sufficient training and test samples, which is becoming more popular especially in the field of automatic driving. The results showed that an effective weighted training strategy for depth prediction were integrated to improve the estimation accuracy. Note that no depth values of image would affect the performances of depth prediction. The proposed new method achieves the state-of-the-art performance on GTA, comparing to DORN. In the future, we will explore further extension on other dense estimation problems.

Acknowledgements this research is supported by Natural Science Foundation of Beijing "Research on the Planning Decision Making Supporting Approaches of Healthy City Planning of Beijing Based on the Analysis of Social Sensing Data" (No. 8182027), and open fund of Institute for China Sustainable Urbanization, Tsinghua University:"Pre-study on new urban development strategy integrating multi-source big data" (TUCSU-K-17026-01). We are also grateful for the computational resources provided by GTA.

References

1. Rupprecht LC, Belagiannis V, Tombari F, Navab N (2016) Deeper depth prediction with fully convolutional residual networks. In: Fourth international conference on 3D Vision (3DV), vol 1, pp 239–248
2. Shelhamer E, Long J, Darrell T (2017) Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal Mach Intell 39(4):640–651
3. Tateno K, Tombari F, Laina I, Navab N (2017) CNN-SLAM: real-time dense monocular SLAM with learned depth prediction. In: IEEE conference on computer vision and pattern recognition, pp 6565–6574
4. Wang P, Shen X, Lin Z, Cohen S, Yuille A (2015) Towards unified depth and semantic prediction from a single image. In: IEEE conference on computer vision and pattern recognition
5. Li B, Dai Y, He M (2018) Monocular depth estimation with hierarchical fusion of dilated CNNs and soft-weighted-sum inference. Pattern Recogn 83:328–339
6. Cao Y, Wu Z, Shen C (2018) Estimating depth from monocular images as classification using deep fully convolutional residual networks. IEEE Trans Circuits Syst Video Technol 28(11):3174–3182
7. Li R, Xian K, Shen C, Cao Z, Lu H, Hang L (2018) Deep attention-based classification network for robust depth prediction. In: IEEE conference on computer vision and pattern recognition (CVPR)

8. Liu F, Shen C, Lin G (2015) Deep convolutional neural fields for depth estimation from a single image. In: IEEE conference on computer vision and pattern recognition
9. Li B, Shen C, Dai Y, Hengel AVD, He M (2015) Depth and surface normal estimation from monocular images using regression on deep features and hierarchical CRFs. In: IEEE conference on computer vision and pattern recognition
10. Dan X, Ricci E, Ouyang W, Wang X, Sebe N (2017) Multi-scale continuous CRFs as sequential deep networks for monocular depth estimation. In: IEEE conference on computer vision and pattern recognition
11. Fu H, Gong M, Wang C, Batmanghelich K, Tao D (2018) Deep ordinal regression network for monocular depth estimation. In: IEEE conference on computer vision and pattern recognition, pp 2002–2011
12. Chen W, Zhao F, Yang D, Jia D (2015) Single-image depth perception in the wild. In: IEEE conference on computer vision and pattern recognition
13. Garg R, Vijay Kumar BG, Carneiro G, Reid I (2016) Unsupervised CNN for single view depth estimation: geometry to the rescue. In: European conference on computer vision
14. Godard C, Aodha OM, Brostow GJ (2017) Unsupervised monocular depth estimation with left-right consistency. In: IEEE conference on computer vision and pattern recognition
15. Kuznietsov GJ, Stückler J, Leibe B (2017) Semi-supervised deep learning for monocular depth map prediction. In: IEEE conference on computer vision and pattern recognition
16. Uhrig J, Schneider N, Schneider L, Franke U, Brox T, Geiger A (2017) Sparsity invariant CNNs. In: 2017 international conference on 3D Vision (3DV), pp 11–20
17. Fischer RP, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: IEEE conference on computer vision and pattern recognition
18. Chollet F, Keras. In: GitHub. https://github.com/fchollet/keras
19. Abadi M et al (2016) TensorFlow: large-scale machine learning on heterogeneous distributed systems. In: Distributed, parallel, and cluster computing
20. Sanner MF (1999) Python: a programming language for software integration and development. J Mol Graph Model 17(1):57–61
21. Meng L (2014) Acceleration method of 3D medical images registration based on compute unified device architecture. Bio-Med Mater Eng 24(1):1109–1116

Vehicle Trajectory Extraction Method Research for Intersection Bayonet Data

Bingjian Yang, Hao Yue, Wencan Gao, Mengyu Zhang and Yang Liu

Abstract In order to extract the vehicle trajectory on the urban road network and provide data basis and technical support for analyzing the balance of supply and demand of urban traffic, the method to extract the vehicle trajectory which is based on intersection bayonet data will be constructed. Firstly, the daily travel trajectory of the vehicle can be obtained by using the data of intersection bayonet and the method of license plate grouping and time order sorting. Secondly, considering the relationship between link travel time and the vehicle to achieve the separation of the vehicle trajectory chain. Finally, the decision indicator of high-grade road proportion will be used to improve the TOPSIS method, and the approach based on the TOPSIS method will be built to reconstruct vehicle trajectory and infer the trajectory of the lost point which is generated by the heuristic depth-first directional algorithm. The paper chooses Suning County in China as an example to verify the reliability of the above approach. Moreover, it is found that the approach can effectively achieve the extraction, separation, and reconstruction of the vehicle trajectory.

Keywords Intersections bayonet data · Trajectory extraction · Heuristic Depth-First directional algorithm · TOPSIS

1 Introduction

Since the vehicle's trajectory in the road network contains rich traffic attributes, the data can be provided for traffic management and planning by acquiring the trajectories of all running vehicles in the road network. In recent years, automatic license plate

B. Yang · H. Yue (✉) · W. Gao · M. Zhang
Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, School of Transportation, Beijing Jiaotong University, Beijing 100044, China
e-mail: yuehao@bjtu.edu.cn

B. Yang
e-mail: 17120913@bjtu.edu.cn

Y. Liu
Nanjing Institute of City & Transport Planning Co., Ltd, Nangjing 210000, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_95

identification data has been used as a new data source for studying urban traffic due to the improvement of license plate recognition accuracy and the coverage rate of bayonets at urban intersections. The intersection bayonet data is a type of automatic license plate recognition data that is named for its collection at the intersection. At present, the relevant researches based on automatic license plate recognition data mainly include the estimation and prediction of the traffic operation status in the road network [1–3], the link travel time calculation and reliability analysis [4, 5], OD matrix estimation [6, 7], vehicle trajectory reconstruction [8–10], and so on. Nowadays, there are a few vehicle trajectories extraction algorithms proposed. For instances, the particle filter model [7, 8, 10] is used to reconstruct the vehicle trajectories in the large-scale urban network; the multi-target decision method [9, 11] is used to reconstruct the incomplete vehicle trajectory. Nevertheless, most of the current studies are limited to high, because of the limitations of the road networks with high bayonets coverage, because of the limitations of the model itself such as calculation complexity, and the testing environment.

This study develops a vehicle trajectory extraction approach for the unique characteristics of the bayonets data of a small city, and that approach draws on RUAN Shubin's vehicle trajectory extraction algorithm. Our contributions include: (1) a link travel time estimation method is proposed to distribute the travel time to links; (2) the trajectory reconstruction, especially the potential vehicle trajectories set generation is re-designed to find the potential paths quickly and accurately. The remainder of the paper is organized as follows, the methodology for trajectory extraction presented in Sect. 2, followed by the Experiment in Sect. 3. Finally, conclusions are drawn.

2 Methodology

This study focuses on developing a method to extract vehicle trajectory using bayonet data. Generally, bayonet data includes the following attributes: vehicle license plate number, time stamp, vehicle license plate color, bayonet location and lane direction in Table 1. The raw trajectory of a vehicle can be extracted merely by connecting the bayonet locations where the vehicle is passing. However, the raw trajectory could be incomplete due to the low bayonet coverage and detection errors. Therefore, a method

Table 1 Bayonet data attributes

Attribute	Description	Sample
VLPN	Vehicle license plate number	"冀 J****X"
VLPC	Vehicle license plate color	Blue
VHPT	The time of the vehicle passing the bayonet	2016/6/1 19:43:25
VLDP	The bayonet location	Qingyuan Street and Zecheng Road
VLLD	The direction of the lane	East to West

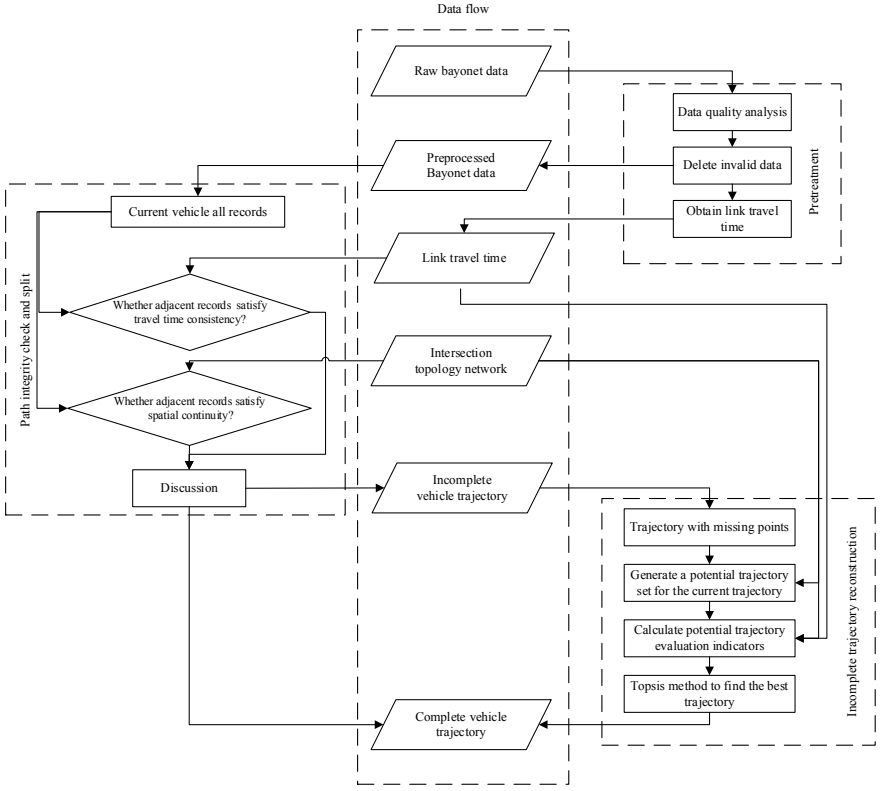


Fig. 1 Flowchart of the proposed method

is required to rebuild complete trajectories firstly, and Fig. 1 shows the flowchart of the proposed method for vehicle trajectories extraction the proposed method consists of three components enclosed by the bold dashed lines in Fig. 1.

(1) Data preparation.

In this component, the data preprocessing includes analyzing the data quality problems existing in the bayonet data, calculating the link travel time between the bayonets, which prepares the data for the separation of the vehicle trajectory and the reconstruction of the incomplete trajectory.

(2) Vehicle trajectory separation.

The trajectory separation is performed by judging the spatial relationship of nodes in the adjacent records and the relationship between the time interval of the next records and comparing the link travel time. It is used to distinguish the complete trajectory from the incomplete trajectory that the spatial relationship between the vehicle passed bayonets.

(3) Vehicle trajectory reconstruction.

In this component, a heuristic depth-first directed search algorithm and a TOPSIS optimization algorithm are established to find the best vehicle travel trajectory from all possible candidate trajectories of each vehicle. The depth-first search algorithm combining the characteristics of the shortest distance heuristic function and the directed search algorithm generates a plurality of candidate trajectories. Secondly, the trajectory decision process calculates various attributes measurements for potential candidate trajectories; finally, the candidate trajectory closest to the real trajectory is selected for rebuilding the incomplete trajectory.

2.1 Data Preparation

It mainly uses three kinds of data: bayonet data, bayonet position data and the link time and space data to realize the extraction and reconstruction of urban vehicle travel trajectories. There are redundant data and error data in the bayonet data, which need to be preprocessed. Meanwhile, the preprocessed bayonet data are calculated to obtain the link travel time.

2.1.1 Bayonet Data Preparation

In the process of acquisition, the bayonet data is affected by many factors such as the accuracy of image recognition, the operation status of the monitoring equipment, and the data submission. The primary data anomalies are shown in Table 2.

This section mainly preprocesses two kinds of data anomaly problems, data duplication and data error, and adopts deletion method for error data. The method of duplicate processing data is as follows:

(1) Eliminate duplicate data

The duplicate data is mainly caused by that the head car behind the signalized intersection stop line may be repeatedly detected during the red light. Therefore, we need

Table 2 Problems of bayonet data

Problems	Problem description
Data missing	Detector missed detection; No detector set at this intersection
Data duplication	Complete same records; Only records whose timestamps differ by a few seconds, and other attributes are identical
Data error	License plate number not recognized; The license plate number does not match the license plate color
Data confusion	The temporal-spatial relationship of the vehicle is disordered

to delete duplicate records from the data set. The following is the method for judging duplicate records:

(a) The same records include the same timestamp.
(b) In addition to timestamp, the adjacent records are the same, and the time interval between two adjacent records is less than the threshold A.

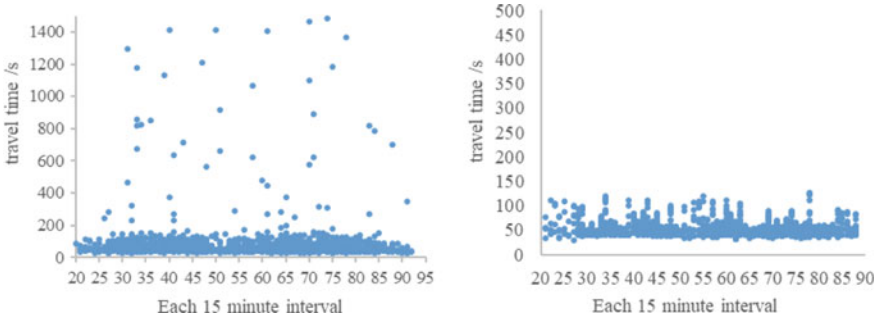
2.1.2 Time and Space Data Acquisition Between Bayonets

The bayonet time and space data includes the travel time and the distance of the adjacent bayonets. Among them, the distance of the adjacent bayonets equals the distance of the link between ones due to the bayonets set intersections. Moreover, the travel time of the adjacent bayonets is obtained as follows: Assume that the time stamps of the adjacent records of a vehicle are a and b, whose bayonets position are m and n. The travel time of the bayonets mn is:

$$t_{\mathrm{mn}} = t_b - t_a \tag{1}$$

The travel time of the adjacent bayonets for a vehicle can be obtained merely by calculating the time delta of the adjacent records. However, the obtained link travel time exiting outlier due to differences in the user's driving behavior such as vehicle pit stop, vehicle over-speed. Therefore, it is necessary to process the outliers of travel time by the method to filter the travel time outliers [11]. Figure 2 shows the changes before and after the link travel time.

If there is no bayonet at the intersection, the vehicles passing through that intersection will not be detected at all. In other words, the intersection will be set a virtual bayonet, which has not set a bayonet and the missing rate of the detected records is 100%. Therefore, the travel time of the link connected to the virtual bayonet cannot be directly obtained. The study found similar characteristics of the bayonet data



(a) Link travel time data before data processing (b) Link travel time data after data processing

Fig. 2 Link travel time data and its outliers

and low-frequency floating car data. We use the method [12] of using low-frequency floating car data to obtain the link travel time, and use the bayonet data to obtain it. We reconstruct the travel time of the link which connected to the virtual bayonet, according to the similarity of other links such as road network topology, socio-economic attributes, and road attributes. We use the weighted interpolation method to distribute the travel time between the two actual bayonets to the link that is connected to the actual bayonet and virtual bayonet. The calculation formula is:

$$t_{ik} = \frac{\alpha l_k}{\sum_k l_k} t_i \tag{2}$$

where: t_{ik} is the travel time of link k in the ith time interval; l_k is the distance of link k; t_i is the travel time between the actual adjacent bayonet; α is the adjustment coefficient, if the angle between adjacent sections less than 45°, $\alpha = 1$; otherwise, $\alpha = 0.9$.

Select 15 min as the statistical time window to obtain the travel time of the link between adjacent bayonets. The link travel time provides data support for separation and reconstruction of vehicle trajectories.

2.2 Vehicle Trajectory Extraction and Separation

This paper defines the vehicle's trajectory as:

$$TR = [point1 \rightarrow point2 \rightarrow \cdots \rightarrow pointn]$$

where *TR* is the travel trajectory chain of the vehicle; *point i* is the location of the bayonet at the time *i* of the vehicle.

The one-day bayonet data is grouped by the license plate number, then sorted in chronological order, the finally the raw trajectory and time stamp of each vehicle are extracted. Since the vehicle has multiple single trips a day, it is necessary to separate the acquired vehicle trajectory into multiple single travel trajectories. Moreover, the travel trajectory is separated by the spatiotemporal relationship of the vehicle motion and the Judging criteria that are the physical property of the bayonet and the time interval in the adjacent records. Standard. The specific vehicle trajectory integrity check and trajectory separated steps are as follows:

Step 1: Generate the topology of the bayonet point, that is, generate the adjacency matrix D of the road intersection, and assume that the intersections a and *b* are connected, then $D(a, b) = 1$; otherwise, $D(a, b) = 1$;

Step 2: The bayonet data is grouped by license plate number and sorted in chronological order;

Step 3: Traverse each license plate number to extract the raw trajectory TR and time stamp of the current vehicle license plate number;

Table 3 Four types of Approach between adjacent nodes of the trajectory chain and their situation

Approach	Situation
Complete trajectory, continue to execute the program	$D(a,b)=1;\ T_{low}<\Delta t<T_{up}$
Incomplete trajectory, execute trajectory reconstruction	$D(a,b)=0;\ T_{low}<\Delta t<T_{up}$
Trajectory separation	$a=b;\ \Delta t>600s$
	$a\neq b;\ \Delta t>T_{up}$
Detection error	$\Delta t<T_{low}$

Step 4: Obtain the bayonet locations a, b and time intervals Δt of the adjacent records of the current vehicle. If $D(a,b)=1$, execute Step 5, otherwise, execute Step 4;

Step 5: Use the shortest path algorithm to obtain the shortest path from a to b. and calculate the upper T_{up} and lower thresholds T_{low} of the shortest path travel times.

Step 6: Discuss the situation and execute the path trajectory extraction as shown in Table 3;

Step 7: Repeat step 4 → step 6.

The method of obtaining T_{up}, T_{low} is as follows:

(1) Design the statistical time window. The traffic flow in the road network is fluctuating within one day. The link travel time obtained during the peak period is relatively longer than that during the peak period. Therefore, one day is subdivided into multiple statistical time windows. Moreover, it considered that the link travel time in the same statistical time window is the same. This paper selects the 5:00–22:00 as the statistical time window of the study, and uses 15 min as the time interval to calculate the link travel time;

(2) Use the travel time acquisition method proposed in the previous section to obtain the link travel time. The use of the filtering method eliminates the travel time that is more discrete, resulting in a reduction in travel time. Therefore, this paper introduces a correction coefficient to correct travel time. The upper and lower thresholds of the link travel time are calculated as follows:

$$T_{up}=\beta\sum_{i}T_{iMax},\ T_{low}=\beta\sum_{i}T_{i\min} \tag{3}$$

where: T_{up}, T_{low} are the corrected upper and lower thresholds of the shortest path travel time; T_{iMax}, T_{iMin} is the upper and lower thresholds of the ith link travel time; β is a correction coefficient which is greater than 1, this paper set $\beta=1.1$.

2.3 Vehicle Trajectory Reconstruction

For the single trip trajectory obtained after the separation of the vehicle's travel trajectory, there may be missing, because the bayonet detection equipment in small cities cannot cover all intersections, and the complete travel trajectory chain requires that the adjacent bayonet is the adjacent intersection. The vehicle trajectory reconstruction method is introduced which apply the TOPSIS method to find the best trajectory as the actual trajectory from the potential trajectory set generated by a heuristic depth-first directed search algorithm.

2.3.1 Potential Trajectories Set Generation

We assume that a single trip will not occur through the same bayonet twice. In order to reduce the computational complexity of the algorithm, a set of potential trajectories for a single trip of the vehicle is created for multi-target decision. The generation process of the potential trajectory is as follows:

Step 1: determine the location of the loose points in the incomplete trajectory, make the bayonet in the previous record of the loose as P1, and the bayonet in the successor record of the loose as P2;

Step 2: The driver's path decision is influenced by many factors, leading to that the actual travel trajectory may not follow the shortest path. Therefore, it is necessary to find multiple alternative paths from P1 to P2, by using a heuristic depth-first directed search algorithm [13] to obtain k candidate paths to form a set of potential travel paths ($TR_1, TR_2, \ldots, TR_k$). In this study, we set k = 10;

Step 3: the successor point (P_{next}) of point P1 and the previous node (P_{pro}) of point P2 are inferred by the road network and the lanes directions of the point P1 and P2,

Step 4: Obtain the corresponding node (P'_{next} , P'_{pro}) of each path candidate, and delete the potential path candidates that are not satisfied $\begin{cases} P_{next} = P'_{next} \\ P_{pro} = P'_{pro} \end{cases}$.

2.3.2 TOPSIS Method Vehicle Travel Track Repair Decision

This study selected the TOPSIS method [14] to make multi-objective decision-making for potential candidate trajectory set, to obtain the best path that best matches the real trajectory. Assume that the potential track set is $TR = [TR_1, TR_2, \ldots, TR_m]$, where $TR_i (i \in M)$ is the ith trajectory candidate.

The driver's path decision is influenced by many factors such as time cost, economic cost, and psychological factors. Among them, the time cost is the travel time of the path candidate; the economic cost is proportional to the distance of the path

candidate; the driver's psychological factors include the number of signal intersections, the number of vehicle turning (especially the number of vehicle left turning) and the proportion of high-grade links on the path.

The decision target set is $X = [X_1, X_2, X_3, X_4, X_5]$, where X_1, X_2, X_3, X_4, X_5 are the distance of the path, the degree of coincidence of the path travel time, the number of intersections, the number of the vehicle turning and the proportion of high-grade links. Due to the data scale of each index is different, this study uses the negative exponent method of Max-Min method to standardize the data. The standardization function is as follows:

(1) The distance of the path: set L_i as the path distance of the ith trajectory candidate, and define the standardized function of the path distance as Formula (4):

$$f_L(TR_i) = \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{L_i - Min_{j\in p}(L_j)}{Max_{j\in p}(L_j) - Min_{j\in p}(L_j)}\right)\right), & Max_{j\in p}(L_j) \neq Min_{j\in p}(L_j) \\ 1, & Max_{j\in p}(L_j) \neq Min_{j\in p}(L_j) \end{cases} \tag{4}$$

(2) The travel time match degree: set t_j as the sum of the travel time averages of the sections of the ith trajectory candidate, and t_* is the actual travel time of path. Define the standardized function of Travel time match degree as Formula (5):

$$f_T(TR_i) = \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{|t_i - t_*| - Min_{j\in p}\left(|t_j - t_*|\right)}{Max_{j\in p}\left(|t_j - t_*|\right) - Min_{j\in p}\left(|t_j - t_*|\right)}\right)\right), & Max_{j\in p}(|t_j - t_*|) \neq Min_{j\in p}(|t_j - t_*|) \\ 1, & Max_{j\in p}(|t_j - t_*|) \neq Min_{j\in p}(|t_j - t_*|) \end{cases} \tag{5}$$

(3) The Number of signalized intersections: Set N_i as the signal intersection number of the ith trajectory candidate, define the standardized function of the signal intersection number as Formula (6):

$$f_N(TR_i) = \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{N_i - Min_{j\in p}(N_j)}{Max_{j\in p}(N_j) - Min_{j\in p}(N_j)}\right)\right), & Max_{j\in p}(N_j) \neq Min_{j\in p}(N_j) \\ 1, & Max_{j\in p}(N_j) \neq Min_{j\in p}(N_j) \end{cases} \tag{6}$$

(4) The Number of the vehicle turnings: Set TN_i as the vehicle turning number of the ith trajectory candidate, and define the standardized function of the vehicle turning number as Formula (7):

$$f_{TN}(TR_i) = \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{TN_i - Min_{j\in p}\left(TN_j\right)}{Max_{j\in p}\left(TN_j\right) - Min_{j\in p}\left(TN_j\right)}\right)\right), & Max_{j\in p}(TN_j) \neq Min_{j\in p}(TN_j) \\ 1, & Max_{j\in p}(TN_j) \neq Min_{j\in p}(TN_j) \end{cases} \tag{7}$$

(5) The proportion of high-grade roads: Set D_i as The vehicle travel distance of the high-grade road of the ith trajectory candidate, and define the standardized function of high-grade roads proportion as Formula (8):

$$f_D(TR_i) = \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{Max_{j\in p}(D_j)-D_i}{Max_{j\in p}(D_j)-Min_{j\in p}(D_j)}\right)\right), & Max_{j\in p}(D_j) \neq Min_{j\in p}(D_j) \\ 1, & Max_{j\in p}(D_j) \neq Min_{j\in p}(D_j) \end{cases} \tag{8}$$

3 Experiment

3.1 Data Description

This paper takes the road network in the downtown area of Suning County as an example. The location of the bayonet points in this area is shown in Fig. 3.

(1) Adopting the bayonet data of the city for one week (May 30, 2016–June 5, 2016), the data includes 1,684,879 vehicle passing records collected by 27 bayonet points in Suning County.

We selected the data mentioned contained about 265,275 records as the experimental data, above on June 1, 2016.

(2) The bayonet point data contains 27 actual bayonet points and nine virtual bayonet points, which are set at 35 intersections. The bayonet points are as shown in Fig. 3.

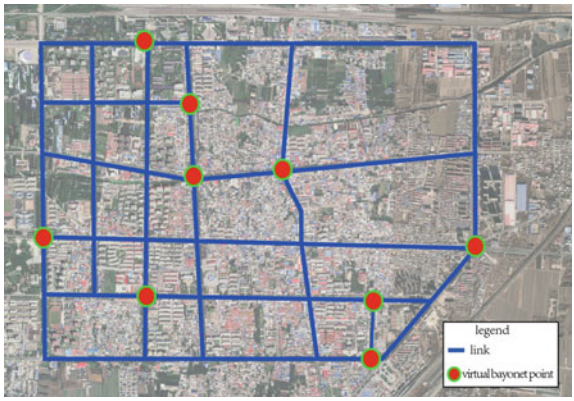


Fig. 3 Bayonet location distribution map

3.2 Example

The current bayonet data set P1 is cleaned to obtain a new data set P2, totaling 211,849 records, and the deleted data was 53,426, accounting for 20.14% of the total data. The deleted data mainly includes abnormal data, duplicate data. The vehicle trajectory is extracted from the data set P2 and constitutes the traveling trajectory set P3 of all the vehicles on the day, after the record of the vehicle trajectory chain length of 1 removed, the remaining number of records is 18,652. After the above steps, useful travel trajectories are extracted as 46,461, and there are 34,631 incomplete trajectories, indicating that the incomplete trajectories situation is abundant. Therefore, the reconstruction of the vehicle trajectory is a vital work. The following is an example to demonstrate the calculation process of the trajectory reconstruction method.

If the trajectory points were lost between node 11 and node 4, the potential trajectory set, TR1 = [11, 7, 8, 9, 3, 4], TR2 = [11, 7, 1, 2, 3, 4], TR3 = [11, 7, 8, 2, 3, 4], is obtained by the heuristic depth-first directed search algorithm, as shown in Fig. 4, the average travel time of each link in each 15 min time interval and the time of the trajectory occurs at 2016-06-01 10:02:21 corresponding to the statistical window 40 in Fig. 5. According to the trajectory of the distance and the spatiotemporal data of the bayonet, five attribute values of each candidate trajectory can be calculated, and the decision matrix is constructed by using the attribute values as follows:

$$P = \begin{matrix} \\ TR_1 \\ TR_2 \\ TR_3 \end{matrix} \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & X_5 \\ 1 & 0.67 & 1 & 0.61 & 0.67 \\ 0.67 & 0.78 & 1 & 0.78 & 1 \\ 0.61 & 1 & 1 & 0.61 & 0.67 \end{bmatrix}$$

(a) Calculate the standardized decision matrix. Standardized method: $r_{ij} = \frac{xij}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} (i = 1, 2, 3, j = 1, 2, 3, 4, 5)$:

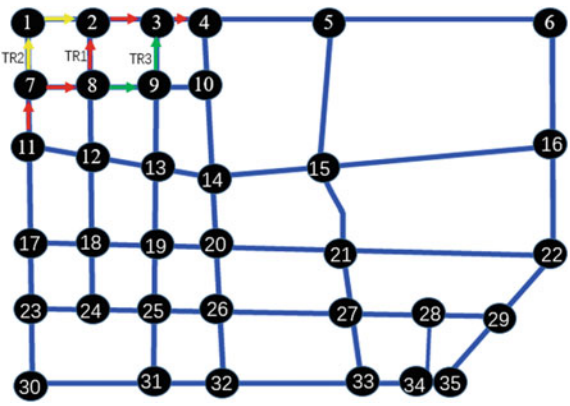


Fig. 4 Three vehicles' trajectories

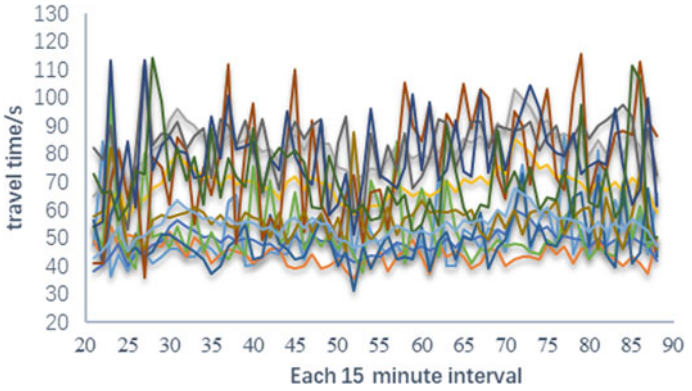


Fig. 5 The average travel time of each link in each time interval

$$R = \begin{matrix} \\ TR_1 \\ TR_2 \\ TR_3 \end{matrix} \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & X_5 \\ 0.74 & 0.47 & 0.58 & 0.52 & 0.49 \\ 0.50 & 0.54 & 0.58 & 0.67 & 0.73 \\ 0.45 & 0.70 & 0.58 & 0.52 & 0.49 \end{bmatrix}$$

(b) Construct a weighted normalization decision matrix. Let the attribute weight vector be $w = [0.2, 0.2, 0.2, 0.2, 0.2]$, and calculate the normalized decision matrix:

$$V = \begin{matrix} \\ TR_1 \\ TR_2 \\ TR_3 \end{matrix} \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & X_5 \\ 0.148 & 0.090 & 0.116 & 0.104 & 0.098 \\ 0.100 & 0.108 & 0.116 & 0.134 & 0.146 \\ 0.090 & 0.140 & 0.116 & 0.104 & 0.098 \end{bmatrix}$$

(c) Determine the positive ideal solutions V^+ and negative ideal solutions V^-:

$$V^+ = \left\{\max V_{ij} | j = 1, 2, 3, 4, 5\right\} = \{0.148, 0.140, 0.116, 0.134, 0.146\}$$

$$V^- = \left\{\min V_{ij} | j = 1, 2, 3, 4, 5\right\} = \{0.090, 0.090, 0.116, 0.104, 0.098\}$$

(d) Calculate the distance between each program and the positive or negative ideal solutions:

$$D_i^+ = \sqrt{\sum_{j=1}^{5} (V_i^+ - V_{ij})^2}, \quad D_1^+ = 0.005704, \quad D_2^+ = 0.003328, \quad D_3^+ = 0.006568$$

$$D_i^- = \sqrt{\sum_{j=1}^{5} (V_i^- - V_{ij})^2}, \quad D_1^- = 0.003364, \quad D_2^- = 0.003628, \quad D_3^- = 0.002500$$

(e) The relative proximity of the calculation to the ideal solution:

$$T_i = \frac{D_i^-}{D_i^+ + D_i^-}, \quad T_1 = 0.371, \quad T_1 = 0.522, \quad T_1 = 0.276$$

(f) According to the size arrangement scheme, TR2 > TR1 > TR3 where TR2 is the best and TR3 is the worst. Therefore, TR2 is selected as the optimal travel trajectory of the vehicle.

4 Conclusion

In this paper, we proposed a vehicle trajectory extraction algorithm based on intersection bayonet identification data. Firstly, we developed a calculation method for the link travel time in the small city, which provides data support for the separation and reconstruction of the vehicle trajectory. Secondly, the separation considering the relationship between link travel time and the vehicle to separate the vehicle trajectory chain into single trip trajectories. Thirdly, the vehicle trajectory reconstruction method fully considers the influence of the driver's decisions that includes time, economic cost and psychological factors. Finally, the data of Suning County urban area is taken as an example to verify the effectiveness of the method.

Acknowledgements This work is supported by the National Natural Science Foundation of China (Gant Nos. 71771013, 51338008, 71621001, 71501011), and the Center of Cooperative Innovation for Beijing Metropolitan Transportation.

References

1. Ju P, Zhou J, Zhang J, Liao R (2014) Analysis of spatio-temporal characteristic in urban traffic flow based on license plate recognition data. Stat Inf Forum 10:66–72

2. Liu C, Li J, Ma L (2017) Traffic state identification method based on card system license plate recognition data. J Qingdao Univ Technol 38(2):90–99
3. Jiang G, Chang A, Niu S (2011) Traffic congestion identification method based on license plate recognition data. J Harbin Inst Technol 43(4):131–135. https://doi.org/10.11918/j.issn.0367-6234.2011.04.026
4. Zhao Z, Ding W, Zhang S (2016) A travel time calculation method based on spatio-temporal data partition of large scale license plate recognition data set. Acta Electron Sin 44(5):1227–1233
5. Gong Y, Luo X, Wang D, Yang S (2018) Urban travel time prediction based on gradient boosting regression tress. J Zhejiang Univ (Eng Sci) 52(3):453–460
6. Ruimin L, Xiyi C, Ruibo Z (2018) OD estimation model based on automatic vehicle identification data. Urban Transp China 16(83)(2):54+87–92
7. Rao WM et al (2018) Origin-destination pattern estimation based on trajectory reconstruction using automatic license plate recognition data. Transp Res Part C-Emerg Technol 95:29–46
8. Feng Y, Sun J, Chen P (2015) Vehicle trajectory reconstruction using automatic vehicle identification and traffic count data. J Adv Transp 49(2):174–194
9. Ruan S, Wang F, Ma D, Jin S, Wang D (2018) Vehicle trajectory extraction algorithm based on license plate recognition data. J Zhejiang Univ (Eng Sci) 52(5):836–844
10. Yang J, Sun J (2015) Vehicle path reconstruction using automatic vehicle identification data: an integrated particle filter and path flow estimator Transport. Res Part C: Emerg Technol 58:107–126
11. Li X, Shi X (2012) Research on the filtering method for travel time outliers. J Wuhan Univ Technol: Transp Sci Eng 36(1):116–119
12. Qu X, Lin C, Yang Z, Shang Q, Cheng Z (2016) Travel time estimation using low-frequency floating car data. J Harbin Inst Technol 48(9):30–34
13. Fang J, Du Z, Zhang F, Zeng Z, Liu R (2013) Heuristic depth-first directional algorithm for shortest path searching in traffic networks. J Zhejiang Univ (Sci Ed) 40(4):469–474
14. Zavadskas EK, Mardani A, Turskis Z et al (2016) Development of TOPSIS method to solve complicated decision-making problems—an overview on developments from 2000 to 2015. Int J Inf Technol Decis Making 15(03):645–682

Critical Section Identification in Road Traffic Network Based on Spatial and Temporal Features of Traffic Flow

Fei Su, Xiaofang Zou, Yong Qin, Shaoyi She and Hang Su

Abstract The capacity of critical section is one of the important reasons for leading to urban road traffic congestion. The identification of critical section has great significance to alleviate traffic congestion, and can provide support for traffic planning, network transformation, residents' travel plans and so on. Based on the spatial and temporal features of traffic flow, the critical section is defined as the one which has more contribution to the overall network and has greater influence on other sections in this paper. In the section importance measurement framework, space-time distribution is used to explain the contribution of one section to the network, space-time influence is measured to describe its influence on other sections, and the critical section is given by the ranking of section importance. At last, the proposed model is applied in a subset of Beijing's road network, and the results show that the model is practical and feasible, and can identify critical section in road traffic network effectively.

Keywords Critical section · Road traffic network · Space-time influence · Space-time distribution · Section importance

F. Su (✉) · S. She · H. Su
China Transport Telecommunications & Information Center,
Beijing 100011, People's Republic of China
e-mail: 11114217@bjtu.edu.cn

National Engineering Laboratory of Transportation Safety & Emergency Informatics,
Beijing 100011, People's Republic of China

F. Su · Y. Qin
School of Traffic and Transportation,
Beijing Jiaotong University, Beijing 100044, People's Republic of China

X. Zou
China Merchants New Intelligence Technology Co., Ltd,
Beijing 100073, People's Republic of China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_96

1 Introduction

With the rapid urban development, the traffic is becoming severe increasingly. Traffic congestion has been heavily fixed eyes on, especially in Beijing, Shanghai and other cities. Traffic congestion usually occurs in some certain areas or sections. The spatial location and capacity of these certain sections often have a great impact on the traffic state of the surroundings. It can be treated as one of the important reasons for leading to road traffic congestion. Therefore, finding the critical or important sections in the road network is an important and critical work in the research on traffic congestion. It has important theoretical significance to alleviate traffic congestion and improve the quality of travel service.

More recently, many researchers have paid attention to the identification of critical section in the road network. In most of which, the importance of section is computed in terms of network topology, connectivity or vulnerability based on complex network theory. Hossain and Sayed [1] gave an introduction to the evaluation index system of consumers' surplus reliability. They measured the importance of road section based on Birnbaum probability. Sanso and Milot [2] analyzed the impact of random events on the reliability of road network. Combined with the traffic state of road network, they proposed that the importance of road section under disaster conditions can be measured according to the degree of impact on road network when it is no longer in force. Nicholson and Du [3] considered that the critical section in a road network is affected by two factors, under the assumption of road capacity changing randomly. One is the contribution of the section to the whole road network; the other is the vulnerability determined by its own attributes. D'Este and Taylor [4] argued that the greater the probability of section being used, the greater the impact on the road network in the process of traffic, and the greater the adverse impact on the road network. Therefore, heuristic algorithm was used to calculate the probability of the impact on the road network after the failure of the section, then the probability was divided by threshold value to find the key section of the network. Tu et al. [5] introduced mincuts frequency vector into communication network technology, and defined the criticality of section in a road network based on the calculation of network topological vulnerability. Comparing the reliability importance of different sections, Hou and Jiang [6] proposed an indirect method to measure the relative importance of sections. On this basis, another indirect method to measure the relative importance of sections was also proposed, and some related examples were given. Taking a small-scale network as an example, Huang [7] solved stochastic user equilibrium assignment model based on the assumption that the capacity obeys a continuous distribution function, then computed section importance using sensitivity analysis, and computed section capacity reliability using capacity reliability distribution function. At last, the ranking of road quality was determined by linear weighted summation. On the basis of complex network theory, Zhao et al. [8] introduced K-shell decomposition method to characterize the location of sections in the road network. They proposed a K-shell based critical section identification method. Furthermore, the influence of section invalid on connectivity reliability of road network was analyzed.

It is turned out that the methodologies work very well for critical section identification. However, it just simply transplanted the complex network theory into transport research, without considering the characteristics of traffic flow. It leads to an inaccurate outcome somewhat. With the deepening of the work, some new methods were applied to critical section identification. Considering traffic flow, section capacity and network topology, Scott et al. [9] used robustness to evaluate the performance of highway network, and compared it with the V/C method. Taylor et al. [10] proposed a method of importance measurement based on link selection probability. It calculated the probability of which section will be chosen from the traveler's view when traffic state was well known. It came come to a decision that the greater the probability, the higher the relative importance of the section. Zou et al. [11] employed a new evaluation method of efficiency of road network using traffic delay. On this basis, the identification of critical section was realized effectively according to efficiency change rate. Based on complex network theory, Liao [12] introduced the concept of proximity and criticality into the algorithm of section importance measurement. Combining with the distribution of traffic flow, the important section in the road network was identified by using the dual theory of graph. Zhong [13] divided the important section into two types: static important section and dynamic important section. Then she gave a definition of important section. Finally, combining with the road network structure and traffic flow characteristics, an evaluation model of important section was established by using connection degree, length, traffic volume, OD pairs and traffic congestion frequency.

In general, there has been an increase in the application of critical section identification. Many fine methodologies have been reported combined with the actual situation of road traffic. However, there is no unified definition of critical section in a road network, as well as a mature theoretical method or evaluation system of critical section identification.

For these reasons, and with the goal of enabling a method to identify the critical section in a road network, the section importance is measured by two aspects in this work: the contribution of a section to the whole network and its influence on other surrounding sections. Thus, the critical section can be defined as the section which contributes more to the traffic flow of the road network and has greater influence on other sections in the network. Firstly, for a section, space-time distribution is proposed to explain its contribution to the traffic flow of the network, and space-time influence is taken to explain its influence on other sections. Then, this paper gives a critical section identification method through sorting the section importance. At last, the proposed model is applied in a subset of Beijing's road network, and the results show that the model is practical and feasible, and can identify critical section in the road network effectively.

The article is structured as follows: the importance of identification of critical section is emphasized in Sect. 1; Sect. 2 gives a description of spatial weight matrices of traffic flow for road networks, to measure the adjacency between any two spatial locations of networks, expressing the spatial layout of the different locations. Section 3 gives a mathematical description and the modeling process of the critical section identification model. For a section, space-time influence analysis is introduced to explain the strength of influence between the section and its neighbors under different time delay, and traffic flow space-time distribution is also analyzed to describe its contribution to the traffic flow of the road network. In Sect. 4, the model has been tested in practice on a subset of Beijing expressway section and the results are analyzed in detail. Finally, we make some conclusions with discussions on future directions in Sect. 5.

2 Spatial Weight Matrix

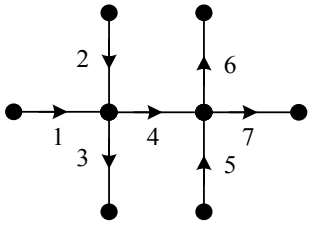
Traffic flow has obvious temporal and spatial characteristics [14]. In order to understand the strength of influence between section and its neighbors in the form of space-time influence, it is necessary to measure the physical relationship between the spatial locations in the road network. Spatial weight matrix is one of the mathematical models to measure the adjacency between any two spatial locations in the network, such as network topological, adjacency relation and so on [15]. A spatial weight matrix W is an $N*N$ matrix, where N is the number of spatial locations. Two issues should be resolved before defining a spatial weight matrix, as different weight matrices can lead to different inferences being drawn [16]. Therefore, this work discusses each of the issues in turn below in the context of road networks.

(1) The structure of networks should be defined to explain which pairs of locations are connected and the direction of connection.
(2) The value of the weights w_{ij} must be defined to describe connection relationship of two locations.

2.1 Road Network Adjacency Matrix

Drawing from graph theory, a network can be viewed as a graph $G = (N, E)$, where N is a set of n nodes, E is a set of edges connecting pairs of nodes. Thus, the network structure can be defined by an $N*N$ adjacency matrix binary [0, 1] [17]. In fact, road network is a complex network composed of sections and intersections, where the intersection represents connection between the sections. Therefore, in this paper, the adjacency matrix can be expressed in the way as below: i and j are edges with variable observations and the nodes represent connections between them. When i and j are directly linked by a node, it can be called first-order neighbors, and the adjacency

(a)



(b)

Segment	1	2	3	4	5	6	7
1	0	0	1	1	0	0	0
2	0	0	1	1	0	0	0
3	1	1	0	0	0	0	0
4	1	1	0	0	0	1	1
5	0	0	0	0	0	1	1
6	0	0	0	1	1	0	0
7	0	0	0	1	1	0	0

(c)

Segment	1	2	3	4	5	6	7
1	0	1	0	0	0	1	1
2	1	0	0	0	0	1	1
3	0	0	0	1	0	0	0
4	0	0	1	0	1	0	0
5	0	0	0	1	0	0	0
6	1	1	0	0	0	0	1
7	1	1	0	0	0	1	0

Fig. 1 A road network and its first and second-order adjacency matrices: **a** the road network, **b** first-order adjacency matrix and **c** second-order adjacency matrix

matrix containing all first-order relations between spatial locations of a network is termed its first-order adjacency matrix W_1. Second-order adjacency matrix W_2 of a network is the first-order matrix of its first-order adjacency matrix and so on. By following ways, adjacency matrix W_k can be defined.

Besides, to define the incidence structure, the direction of influence should be also considered. In our opinion, road network can be viewed as a directed graph when considering the traffic flow [18]. Furthermore, the adjacency matrix of road network is different from other directed networks:

(1) The traffic only flows from upstream to downstream, but the influence may be occurred in both directions. In free conditions, the influence will be from upstream to downstream, while in congested conditions, it will be mainly from downstream.
(2) When two sections flowing into or out of the same link from the same direction, the influence may be mostly from the second-order instead of first-order neighbors.

As the road network shown in Fig. 1, it is directed, adjacency matrix is asymmetric since the influence goes in one direction and an edge only influences its edges downstream. In the network, the traffic of section 1 and 2 flow into section 4. The increase in traffic of section 1 will lead to the congestion of section 4, then downstream situation of section 4 feedback an influence on section 1. Meanwhile, the congestion will also propagate up to section 2. In a word, sections 1 and 4 has a direct effect between each other, while section 1 and 2 has an indirect effect via section 4. In other words, sections 1 and 2 can be treated as second-order neighbors, while both as first-order neighbors of section 4.

In summary, the adjacency matrix of road networks can be defined according to the rules above in the case study. Figure 1 gives the example of the first and second-order adjacency matrix of the road network.

2.2 Adjacency Matrix Weights

Many rules have been used in the literature to Defining the weights w_{ij} of the spatial weight matrix. Usually, it is denoted by binary 0 and 1, and if the relationship exists between i and j, it will be 1. Otherwise, it will be 0. For example, there lists two commonly used ways as follows [19, 20].

(1) Simple adjacency relationship

$$w_{ij} = \begin{cases} 1, \text{ when } i \text{ and } j \text{ has connection relationship} \\ 0, \text{ else} \end{cases} \tag{1}$$

(2) Adjacency relationship based on distance

$$w_{ij} = \begin{cases} 1, \text{ when the distance between } i \text{ and } j \text{ is less than } d, \\ 0, \text{ else} \end{cases} \tag{2}$$

In practice, the definition of adjacency matrix weights is very much dependent on the type of data under study, for example, simple adjacency may be used for areal data, while distance adjacency or nearest-neighbors may be used for point data.

In this work, the kth-order adjacency matrix weights are denoted as Eq. (3), considering the distance between two sections, to explain the road network structure. Taking the road network shown in Fig. 1 as an example, the spatial weight matrix can be obtained as in Fig. 2 through the definition of adjacency matrix and spatial weights.

$$w_{ij} = \begin{cases} \frac{1}{k}, \text{ when } i \text{ and } j \text{ are } k\text{th-order neighbors} \\ 0, \text{ else} \end{cases} \tag{3}$$

(a)

Segment	1	2	3	4	5	6	7
1	0	0	1	1	0	0	0
2	0	0	1	1	0	0	0
3	1	1	0	0	0	0	0
4	1	1	0	0	0	1	1
5	0	0	0	0	0	1	1
6	0	0	0	1	1	0	0
7	0	0	0	1	1	0	0

(b)

Segment	1	2	3	4	5	6	7
1	0	1/2	0	0	0	1/2	1/2
2	1/2	0	0	0	0	1/2	1/2
3	0	0	0	1	0	0	0
4	0	0	1/2	0	1/2	0	0
5	0	0	0	1	0	0	0
6	1/2	1/2	0	0	0	0	1/2
7	1/2	1/2	0	0	0	1/2	0

Fig. 2 Spatial weight matrices of road network: **a** first-order spatial weight matrix; **b** second-order spatial weight matrix

3 Critical Section Identification Model

Based on the temporal and spatial features of traffic flow, the modeling process of the critical section identification is introduced. In this work, the importance of section is measured in terms of its traffic flow contribution to the network and the influence on other sections at the same time. For this, the critical section is identified by sorting the section importance, and it can be defined as the section which contributes more to the traffic flow of the road network and has greater influence on other sections in the road network. Then, space-time influence is measured using the cross-correlation function (CCF) to compute the strength of influence between section and its neighbors; Space-time distribution is analyzed to explain the traffic flow contribution of the section. Therefore, the two issues are discussed in turn below. At last, an evaluation model of section importance is built through the linear weighted on the two metrics, accordingly, to identify the critical section in a road network by sorting.

3.1 Space-Time Influence

In 2008, Box et al. [21] employed the cross-correlation function to measure the correlation of two spatial objects at a given delayed time. As a correlation measure, the CCF treats two traffic time series as a bivariate stochastic process and measures correlation of traffic flow between a certain section in the road network and its neighbors at s time lags. Given two time series X and Y, the CCF at s time lags can be denoted as follows:

$$\gamma(s) = \frac{E(x(t) - \mu_X)(y(t+s) - \mu_Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}} \tag{4}$$

where $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ are the means and variances of time series X and Y respectively. $x(t)$ is the observations of X at time t, and $y(t+s)$ is the observations of Y at time $t+s$.

However, there are usually not only one kth-order neighbor of a certain section in the network. For this, the spatial weight matrix is drawn into the CCF, to represent the space-time influence between a section and its kth-order neighbors clearly and accurately. On the basis of it, given the traffic flow time series X of a section and the time series Y weighted its kth-order spatial neighbors at s time lags, the CCF can be developed as follows:

$$\gamma_i(k, s) = \frac{\sum_{t=1}^{T} (x_i(t) - \overline{x_i})\left(W^{(k)} x_i(t+s) - \overline{W^{(k)} x_i}\right)}{\sqrt{\sum_t (x_i(t) - \overline{x_i})^2} \sqrt{\sum_t \left(W^{(k)} x_i(t+s) - \overline{W^{(k)} x_i}\right)^2}} \tag{5}$$

where $\gamma_i(k, s)$ is the space-time correlation between section i and its kth-order neighbors at time lag s; T is the period of statistical time; $x_i(t)$ is traffic flow observations of the given section i at time t; $\overline{x_i}$ is the average value of time series X; $W^{(k)}x_i(t+s)$ is a space delay operator, representing the weighted mean traffic observations of all kth-order neighbors of the section i at time lag s, and $\overline{W^{(k)}x_i}$ is the average value of the kth-order neighbors' traffic, then it can be denoted as follows:

$$W^{(0)}x_i(t) = x_i(t), \quad W^{(k)}x_i(t) = \sum_{j=1}^{m} w_{ij}^{k} x_j(t) \tag{6}$$

where m is the number of the section in the network, w_{ij}^{k} is the weight between section i and j in the kth-order spatial weight matrix.

Space-time correlation can be employed to understand the strength and direction of the effect between a section and its kth-order neighbors [22]. For example, traffic flows downstream from upstream with some time lags, and the traffic situation of upstream will influence on downstream, meanwhile, it will also feedback an effect on upstream from downstream due to the congestion. In this paper, space-time correlation can be used to explain the relationship of the influence as below.

- If $\gamma_i(k, s)$ is significantly different from 0 and $s > 0$, it shows that the current traffic state of section i has influence on the future state of the kth-order neighbors of section i.
- If $\gamma_i(k, s)$ is significantly different from 0 and $s < 0$, it means that the current traffic state of the kth-order neighbors of section i has influence on the future state of section i.
- If $\gamma_i(k, s)$ is significantly different from 0 and $s = 0$, it indicates that the traffic state is very similarly at the same time between section i and its neighbors, and it does not capture the direction of influence of one section on another.

Note that space-time correlation measures the influence of one section on its kth-order neighbors under a time lag s, and it can be treated as a discrete process varying with time delay s and space delay k. In order to calculate the influence of a section on other sections around in the network comprehensively, it is necessary to synthesize space-time influences in spatial and temporal dimensions respectively. Therefore, an integrated spatial weight matrix is considered in spatial dimension. It can be obtained by the sum of different order spatial weight matrix as the following equation:

$$W' = \sum_{i=1}^{k} W_i \tag{7}$$

where W_i is the ith-order spatial weight matrix. W' is the integrated spatial weight matrix, representing that all indirect or direct influence of one section on its neighbors from first to kth-order will be considered simultaneously. It is worth noting that the weights w'_{ij} should be row standardized in the case study.

In addition, the technique for order preference by similarity to an ideal solution (TOPSIS) is employed as a synthetical method for space-time correlation in temporal dimension to obtain the space-time influence. Its steps are as follows [23].

(1) Defining positive and negative ideal points

$$A^+ = \left\{ \max_i r_i(s) | s \in \mathbf{S}, 1 \le i \le N \right\} = \left\{ A^+(s) | s \in \mathbf{S} \right\},$$
$$A^- = \left\{ \min_i r_i(s) | s \in \mathbf{S}, 1 \le i \le N \right\} = \left\{ A^-(s) | s \in \mathbf{S} \right\} \tag{8}$$

where $r_i(s)$ is space-time correlation of section i at time lag s synthesized in spatial dimension. S is the set of time delay values. N is the number of sections in the road network. $A^+(s)$ is the maximum value of space-time influences. $A^-(s)$ is the minimum value of space-time influences. In fact, it is usually carried out that $A^+(s) = 1, A^-(s) = -1$, based on the range of $[-1, 1]$ of space-time correlation.

(2) Computing distance
The Euclidean weighted distance is usually used as follows to calculate the distance between space-time correlation and the positive and negative ideal points.

$$E_i^+ = \sqrt{\sum_s w_s (r_i(s) - A^+(s))^2}, \quad 1 \le i \le N,$$
$$E_i^- = \sqrt{\sum_s w_s (r_i(s) - A^-(s))^2}, \quad 1 \le i \le N \tag{9}$$

where w_s is weighting coefficient. E_i^+ is the distance between space-time correlation and the positive ideal point. E_i^- is the distance between space-time correlation and the negative ideal point.

(3) Calculating the influence degree
The relative closeness e_i of space-time correlation of section i to the ideal point is calculated to measure the influence degree of section i on other sections in the road network.

$$e_i = \frac{E_i^-}{E_i^- + E_i^+}, \quad 1 \le i \le N \tag{10}$$

(4) Normalization of the influence degree
So as to distinguish the differences of the influence of a section on others easily, the influence degree should be normalized as follows.

$$c_i = \frac{e_i - \min e_i}{\max e_i - \min e_i}, \quad 1 \le i \le N \tag{11}$$

where c_i is the space-time influence of the section i.

3.2 Space-Time Distribution

In this paper, the section importance is measured by two aspects based on the spatial and temporal features of traffic flow: space-time influence and space-time distribution. In which, space-time distribution can be explained as how much the traffic state of a section contributes to the whole road network. Traffic flow is changing with time and space, the traffic change of the section is analyzed using the statistical method, to understand the distribution and evolution of traffic flow in the road network. Then, space-time distribution can be calculated to achieve the critical section identification. For example, at a certain moment, the heavier the traffic congestion of a section (that is, the greater the traffic volume, or the longer the travel time, or the lower the travel speed of the section), the greater the contribution of the section to the road network traffic state, which also reflects the greater importance of the section in the road network [24]. Meanwhile, traffic flow changes periodically, and it changes with time in each cycle. Therefore, the traffic flow observation at every moment in a cycle should be considered, when computing the space-time distribution. On the basis of this, the distribution equilibrium of traffic flow in the road network can be further studied in the future.

Given the traffic flow observations of a road network, the space-time distribution can be denoted as follows:

$$d_i = \frac{s_i - \min s_i}{\max s_i - \min s_i}, \quad 1 \le i \le N \tag{12}$$

where d_i is the space-time distribution of section i. s_i expresses the contribution of the section i to the road network in a time period T. $\max s_i$ is the maximum value of all sections' contributions to the road network. $\min s_i$ is the minimum value of all sections' contributions to the road network. N is the number of sections in the road network.

$$s_i = \sum_{t=1}^{T} \frac{\theta_x^T(i,t)x_i(t)}{w_x^N(t)}, \theta_x^T(i,t) = \frac{x_i(t)}{\sum_{t=1}^{T} x_i(t)}, w_x^N(t) = \sum_{i=1}^{N} x_i(t) \tag{13}$$

where $x_i(t)$ is the traffic flow observation of section i a certain time t. $\theta_x^T(i,t)$ is the traffic flow temporal distribution of section i a certain time t in the statistical time period. $w_x^N(t)$ is the sum of real-time traffic flow observations in all sections of road network at a certain time t. T is the statistical time period.

3.3 Importance Calculation

In this work, the section importance is measured by two aspects in this work: the contribution of a section to the whole network and its influence on other surrounding sections. The critical section is identified by sorting the section importance, and it can be defined as the section which contributes more to the traffic flow of the road network and has greater influence on other sections in the road network. Thus, space-time influence and space-time distribution should be considered at the same time to calculate the importance of the section. In order to get the comprehensive value for the importance of a section in the road network, the linear weighting method is employed, and it can be denoted as follows:

$$I_i = \alpha \cdot c_i + (1 - \alpha) \cdot d_i \tag{14}$$

where I_i is the importance measurement of the section i. α is the weight, determining the key factor in measuring the section importance. When the weight α is large ($\alpha > 0.5$), it means that more attention will be paid to space-time influence of the section on its surroundings; When the weight α is small ($\alpha < 0.5$), it means that more consideration will be paid to space-time contribution of the section to the road network.

4 Case Study

In the case study, the measurement of section importance is applied to traffic data in practice on Beijing's road network in order to identify the critical section. Beijing Traffic Management Bureau has built the Expressway Traffic Information Detection System. It collects real-time volume, speed and time occupancy for expressways depending on the microwave detectors. These sensors cover nearly all the expressways including ring roads and connecting expressways. Although the incidence structure of the sensor network does not fully mimic that of the real Beijing's road network. So many junctions may be included in one section and many minor roads are omitted due to the limit of setting up the sensors. A subset of Beijing's road network is chosen as test network in the case study. As mentioned in reference to Fig. 3 and Table 1, the test network comprises 26 sections.

The traffic data were obtained daily from the detectors every 2 min. For the survey data sets, because the 2-min data collected by the sensors are all realistic data and there is too much noise in the data, the modeling is performed on traffic flow data transformed into 10 min discrete time intervals. That is, there are 144 observations in one day and it means 10 min for one-time interval. Finally, traffic flow data for six weeks from 9 May to 19 June 2011 is selected in practice. Furthermore, the smooth operation is necessary to significantly reduce the random elements caused by mixed traffic and other factors.

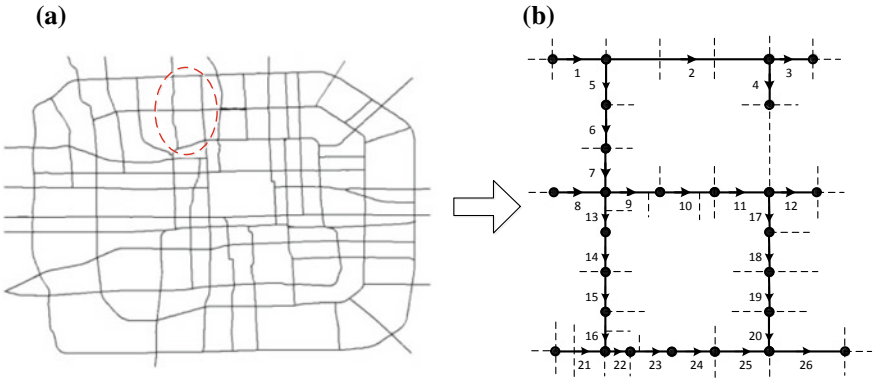


Fig. 3 A subset of Beijing's road network as the test network: **a** the location of the test network, **b** the section in the test network

Table 1 The origin and destination of the 26 sections in the test network

Section	Origin	Destination	Section	Origin	Destination
1	Zhan Chun Second Bridge	Xue Yuan Bridge	14	Ming Guang Bridge North	Ming Guang Bridge
2	Xue Yuan Bridge	Jian Xiang Bridge	15	Ming Guang Bridge	Wen Hui Bridge
3	Jian Xiang Bridge	Bei Chen Bridge West	16	Wen Hui Bridge	Xi Zhi Men Bridge
4	Jian Xiang Bridge	Jian Xiang Bridge North	17	Ma Dian Bridge	Bei Jiao Market
5	Xue Yuan Bridge	Hua Yuan North Road	18	Bei Jiao Market	Xin Kang Road
6	Hua Yuan North Road	Xue Zhi Bridge	19	Xin Kang Road	An De Road
7	Xue Zhi Bridge	Ji Men Bridge	20	An De Road	De Sheng Men Bridge
8	Capital University of Physical Education and Sports	Ji Men Bridge	21	Bei Zhan Bridge	Xi Zhi Men Bridge
9	Ji Men Bridge	Hua Yuan Road	22	Xi Zhi Men Bridge	Huapi Factory Hutong
10	Hua Yuan Road	Bei Tai Ping Zhuang Bridge	23	Huapi Factory Hutong	Xin Jie Kou
11	Bei Tai Ping Zhuang Bridge	Ma Dian Bridge	24	Xin Jie Kou	Ji Shui Tan Bridge

(continued)

Table 1 (continued)

Section	Origin	Destination	Section	Origin	Destination
12	Ma Dian Bridge	Yu Min Middle Road	25	Ji Shui Tan Bridge	De Sheng Men Bridge
13	Ji Men Bridge	Ming Guang Bridge North	26	De Sheng Men Bridge	Drum Tower Bridge

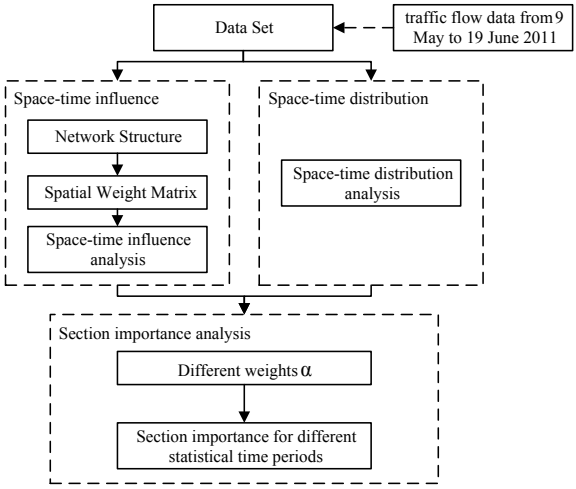


Fig. 4 Process of the case study

In order to identify the critical section throughout the day in detail, the modeling is performed according to the three statistical time periods: 0:00–24:00, AM peak (7:00–10:00) and PM peak (17:00–20:00). It has been widely accepted in transport studies that traffic behaves differently in each time period.

In this study, as shown in Fig. 4, the analysis is threefold. Firstly, space-time influences of each section in the test network are calculated for the six weeks from 9 May to 19 June 2011 in practice. Secondly, space-time distributions of each section are computed for the 42 days. Finally, we focus on the section importance calculation under the different weights α to analyze the changes of critical section in the road network.

(1) space-time influence

Due to the length of the sections, the neighbors of a certain section higher than five orders may be outside one-time lag (it means 10 min in this case study). Therefore, in this case study, five spatial weight matrices (W_1–W_5) are employed to obtain the integrated spatial weight matrix W' according to Eq. (7). Tables 2 and 3 give the example of the third-order spatial adjacency matrix W_3 and the integrated spatial weight matrix W' of the test network, respectively. Note that the weights should be row standardized in the case study.

Table 2 The third-order spatial weight matrix W3 of the test network

Section	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	1/3	1/3	0	0	0	0	1/3	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	1/3	0	0	0	0	0	1/3	0	1/3	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0
7	1/3	0	0	0	0	0	0	0	0	0	1/3	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	1/3	0	0	0	0	1/3	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	1/3	0	0	0	0	0	0	1/3	0	1/3	0	0	1/3	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	1/3	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	1/3	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0
13	0	0	0	0	1/3	0	0	0	0	1/3	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	1/3	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0
15	0	0	0	0	0	0	1/3	1/3	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	1/3	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	1/3	0	0
17	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	1/3	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3
19	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	1/3	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	1/3	0	0

(continued)

Table 2 (continued)

Section	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
22	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	1/3	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0	0	0	0	0	0	0	1/3
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	1/3	1/3	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	1/3	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/3	0	0	0	0	1/3	0	0	0

Table 3 The integrated spatial weight matrix W' of the test network

Section	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	1	0.5	0.5	1	0.5	1/3	0.2	0.25	0.2	0	0	0.25
2	1	0	1	1	0.5	1/3	0.25	0	0.2	0	0	0	0.2
3	0.5	1	0	0.5	1/3	0.25	0.2	0	0	0	0	0	0
4	0.5	1	0.5	0	1/3	0.25	0.2	0	0	0	0	0	0
5	1	0.5	1/3	1/3	0	1	0.5	0.25	1/3	0.25	0.2	0	1/3
6	0.5	1/3	0.25	0.25	1	0	1	1/3	0.5	1/3	0.25	0.2	0.5
7	1/3	0.25	0.2	0.2	0.5	1	0	0.5	1	0.5	1/3	0.25	1
8	0.2	0	0	0	0.25	1/3	0.5	0	1	0.5	1/3	0.25	1
9	0.25	0.2	0	0	1/3	0.5	1	1	0	1	0.5	1/3	0.5
10	0.2	0	0	0	0.25	1/3	0.5	0.5	1	0	1	0.5	1/3
11	0	0	0	0	0.2	0.25	1/3	1/3	0.5	1	0	1	0.25
12	0	0	0	0	0	0.2	0.25	0.25	1/3	0.5	1	0	0.2
13	0.25	0.2	0	0	1/3	0.5	1	1	0.5	1/3	0.25	0.2	0
14	0.2	0	0	0	0.25	1/3	0.5	0.5	1/3	0.25	0.2	0	1
15	0	0	0	0	0.2	0.25	1/3	1/3	0.25	0.2	0	0	0.5
16	0	0	0	0	0	0.2	0.25	0.25	0.2	0	0	0	1/3
17	0	0	0	0	0	0.2	0.25	0.25	1/3	0.5	1	0.5	0.2
18	0	0	0	0	0	0	0.2	0.2	0.25	1/3	0.5	1/3	0
19	0	0	0	0	0	0	0	0	0.2	0.25	1/3	0.25	0
20	0	0	0	0	0	0	0	0	0	0.2	0.25	0.2	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0.2
22	0	0	0	0	0	0	0.2	0.2	0	0	0	0	0.25

(continued)

Table 3 (continued)

Section	1	2	3	4	5	6	7	8	9	10	11	12	13
23	0	0	0	0	0	0	0	0	0	0	0	0	0.2
24	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0.2	0	0
Section	14	15	16	17	18	19	20	21	22	23	24	25	26
1	0.2	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0.25	0.2	0	0	0	0	0	0	0	0	0	0	0
6	1/3	0.25	0.2	0.2	0	0	0	0	0	0	0	0	0
7	0.5	1/3	0.25	0.25	0.2	0	0	0	0.2	0	0	0	0
8	0.5	1/3	0.25	0.25	0.2	0	0	0	0.2	0	0	0	0
9	1/3	0.25	0.2	1/3	0.25	0.2	0	0	0	0	0	0	0
10	0.25	0.2	0	0.5	0	0.25	0.2	0	0	0	0	0	0
11	0.2	0	0	1	0.5	0	0.25	0	0	0	0	0	0.2
12	0	0	0	0.5	1/3	0.25	0.2	0	0	0	0	0	0
13	1	0.5	1/3	0.2	0	0	0	0.2	0.25	0.2	0	0	0
14	0	1	0.5	0	0	0	0	0.25	1/3	0.25	0.2	0	0
15	1	0	1	0	0	0	0	1/3	0.5	1/3	0.25	0.2	0
16	0.5	1	0	0	0	0	0	0.5	1	0.5	1/3	0.25	0.2
17	0	0	0	0	1	0.5	1/3	0	0	0	0	0.2	0.25

(continued)

Table 3 (continued)

Section	14	15	16	17	18	19	20	21	22	23	24	25	26
18	0	0	0	1	0	1	0.5	0	0	0	0.2	0.25	1/3
19	0	0	0	0.5	1	0	1	0	0	0.2	0.25	1/3	0.5
20	0	0	0	1/3	0.5	1	0	0	0.2	0.25	1/3	0.5	1
21	0.25	1/3	0.5	0	0	0	0	0	1	0.5	1/3	0.25	0.2
22	1/3	0.5	1	0	0	0	0.2	1	0	1	0.5	1/3	0.25
23	0.25	1/3	0.5	0	0	0.2	0.25	0.5	1	0	1	0.5	1/3
24	0.2	0.25	1/3	0	0.2	0.25	1/3	1/3	0.5	1	0	1	0.5
25	0	0.2	0.25	0.2	0.25	1/3	0.5	0.25	1/3	0.5	1	0	1
26	0	0	0.2	0.25	1/3	0.5	1	0.2	0.25	1/3	0.5	1	0

Figures 5, 6 and 7 show the example of the space-time correlation for nine sections in the test road network. The vertical axis shows the value of the space-time correlations, and the horizontal axis shows the time lags, where the horizontal axis unit is 10 min. It is worth noting that it is also performed according to traffic data of three periods: 0:00–24:00, AM peak (7:00–10:00), and PM peak (17:00–20:00).

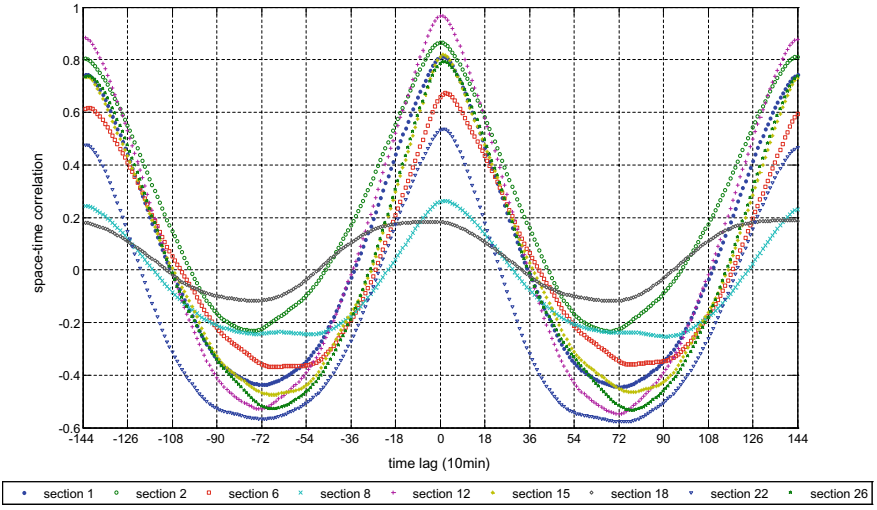


Fig. 5 Space-time correlation of 9 sections in the test road network measured from 0:00 to 24:00

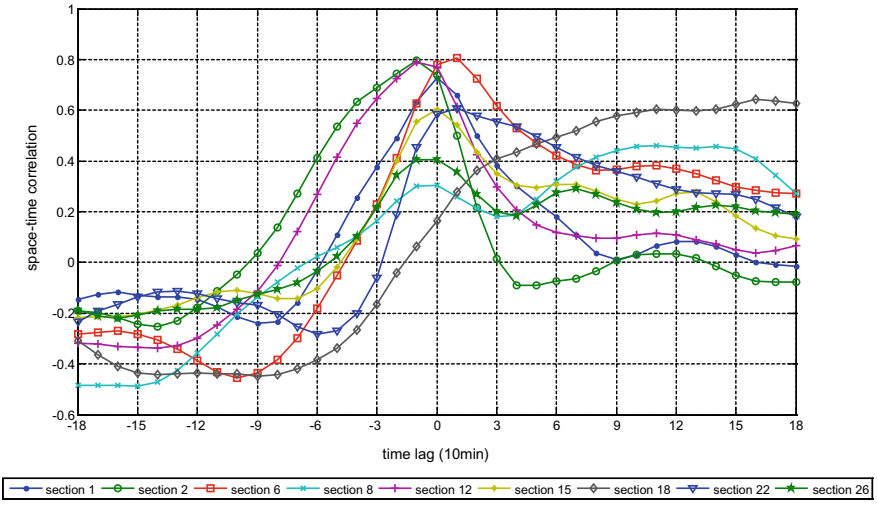


Fig. 6 Space-time correlation of 9 sections in the test road network for the AM peak

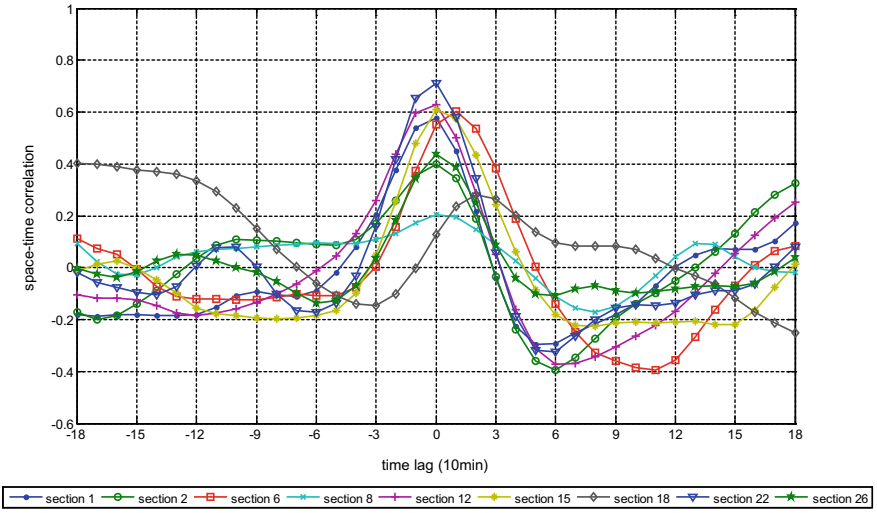


Fig. 7 Space-time correlation of 9 sections in the test road network for the PM peak

It is clear that there is obvious heterogeneity in the space-time correlations, that is, there is difference in the strength of influence on the surroundings for each section. For example, the correlation is obviously different between section 1 and 18 in the statistical period of 0:00–24:00. Meanwhile, for the same section, there is also some difference in different statistical periods. It is that the influence of section on its surrounding sections varies with time. For example, the space-time correlation of section 18 changes from negative to positive with time delay in AM peak, while it does change differently in PM peak.

On the basis of TOPSIS in Sect. 3.1, Fig. 8 gives the values of space-time influence for the 26 sections in the subset of Beijing's road network, performed according to 0:00–24:00, AM peak and PM peak, respectively. Examining the figure, the following statements could be drawn:

- It is different in the strength of space-time influence on others for each section. In this case study, it is 0.66, 0.25 and 0.34 for section 1 in the statistical periods of 0:00–24:00, AM peak and PM peak, respectively, while it is 0.97, 0.26 and 0.53 for section 2.
- Due to the spatial and temporal characteristics of traffic flow, Fig. 8 shows that section 4 has the greatest influence on its surroundings in the period of 0:00–24:00, however, it is section 20 in AM peak, and section 18 in PM peak.

(2) Space-time distribution

According to Sect. 3.2, for the three periods of 0:00–24:00, AM peak and PM peak, Fig. 9 shows space-time distribution of the 26 sections in the test road network, respectively, to measure how much contribution the section makes to the overall network. Examining the figure, the following statements could be drawn:

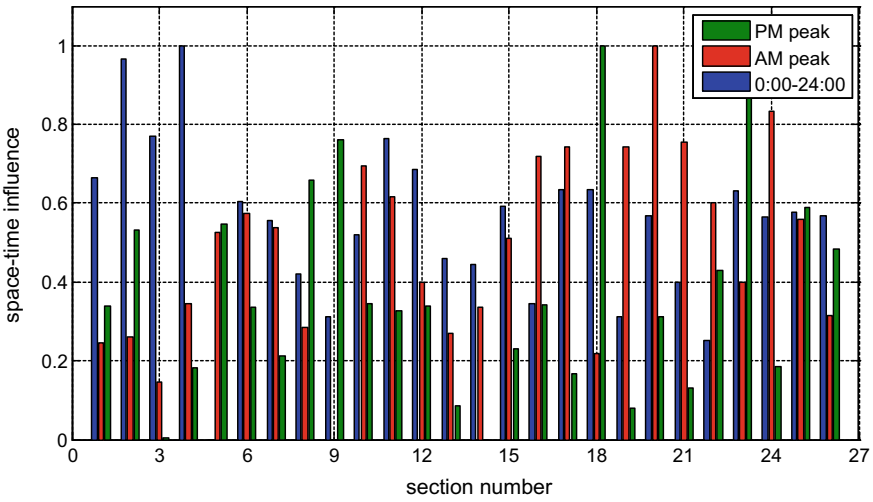


Fig. 8 The day average space-time influence of the 26 sections performed by 0:00–24:00, AM peak and PM peak

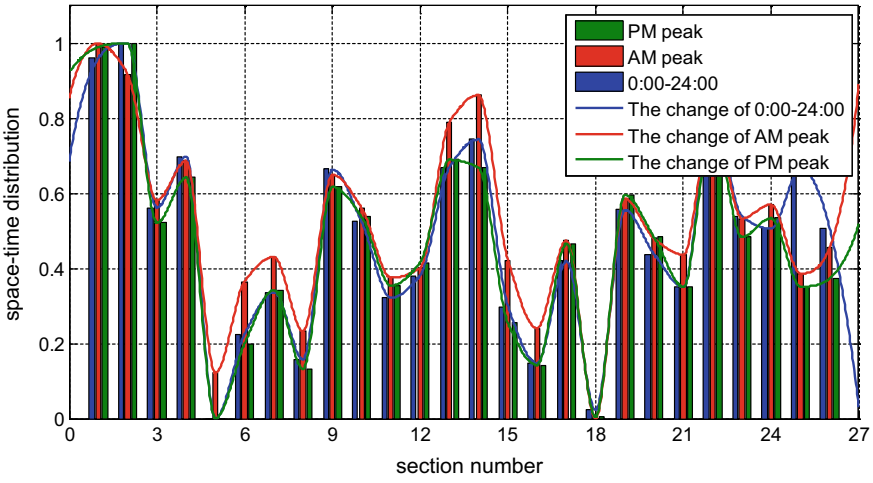


Fig. 9 The day average space-time distribution of the 26 sections performed by 0:00–24:00, AM peak and PM peak

- Figure 9 displays a comparison of space-time distribution of the 26 sections. It could be observed that the values of space-time distribution are obviously different from each other for the three statistical periods, but the change trends of space-time distribution during the three periods are basically the same.

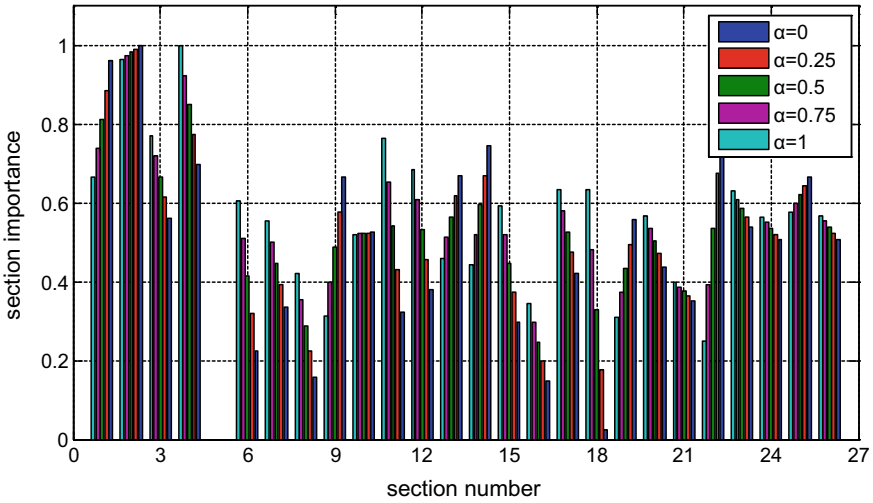


Fig. 10 The section importance under different weights for the period of 0:00–24:00

- In this case study, it is clear that section 2 makes the greatest contribution to the test network during 0:00–24:00, while, it is section 1 in AM peak, and section 2 in PM peak.

(3) Section importance analysis

In order to identify the critical section, space-time influence and space-time distribution are both considered as the measurement of the section importance. According to the linear weighting method mentioned above, Figs. 10, 11, 12 give the comprehensive value for the importance of the 26 sections under different weights, for the three periods of 0:00–24:00, AM peak and PM peak. In this case study, the weights α of 0, 0.25, 0.5, 0.75 and 1 are selected for experimental analysis. Meanwhile, take $\alpha = 0.25$ as an example, Table 4 lists the measurements of section importance and their rankings in importance correspond to 0:00–24:00, AM peak and PM peak.

Examining these figures and table, the following statements could be drawn:

- The weight α determines which factor is considered more in measuring the section importance. For example, Fig. 10 shows the example of the 26 sections importance under different weights for the period of 0:00–24:00. In this case study, the importance of section 1 is 0.9607, 0.8867, 0.8127, 0.7387, 0.6648 corresponding to $\alpha = 0, 0.25, 0.5, 0.75, 1$. However, due to the spatial heterogeneity of traffic flow, the section importance of section 18 is 0.0226, 0.1758, 0.3290, 0.4822 and 0.6353, when the value of weights is 0, 0.25, 0.5, 0.75 and 1 respectively.
- In this case study, it is can be seen that section 2 is the most critical section in the test road network for the period of 0:00–24:00 when $\alpha = 0.25$, and its section importance is 0.9914. While it is section 1 in AM peak (The section importance is 0.8114), and section 2 in PM peak (The section importance is 0.8833).

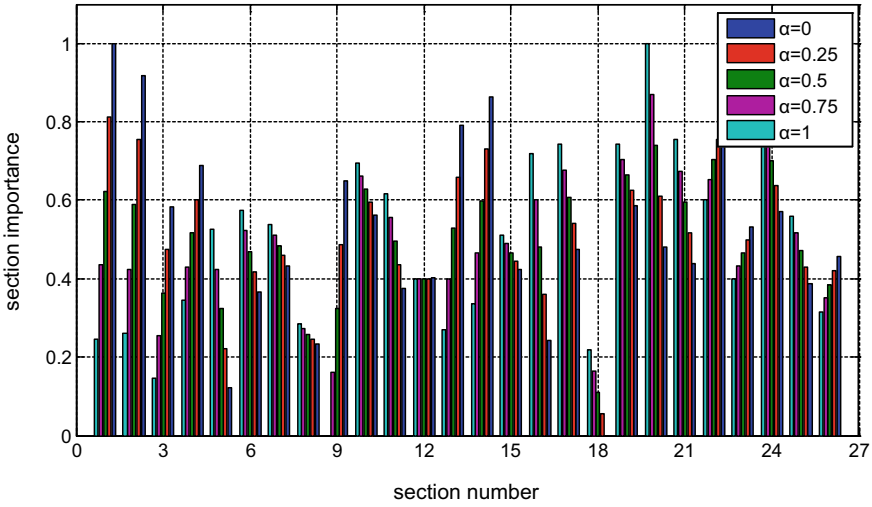


Fig. 11 The section importance under different weights for the AM peak

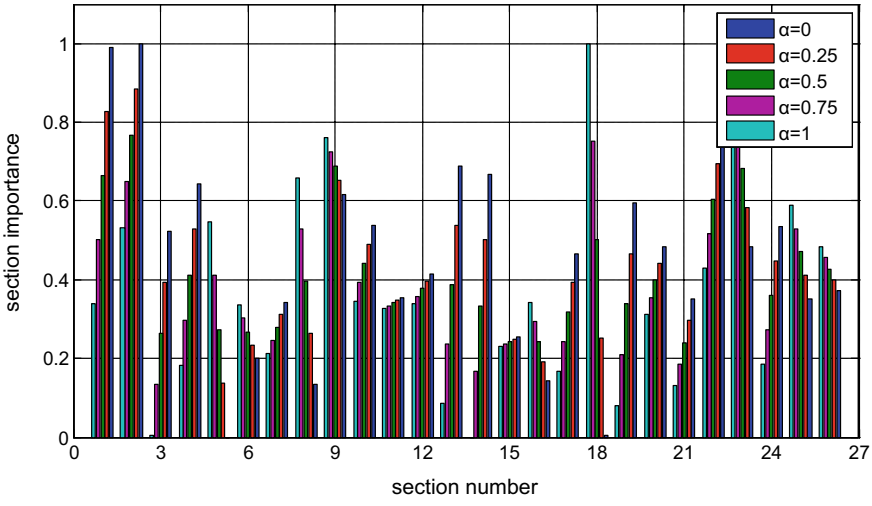


Fig. 12 The section importance under different weights for the PM peak

- When $\alpha = 0.5$, section 2 is the most critical section from 0:00 to 24:00 (The section importance is 0.9828), while it is section 20 in AM peak (The section importance is 0.7401), and section 2 in PM peak (The section importance is 0.7666).
- When $\alpha = 0.75$, section 2 is the most critical section from 0:00 to 24:00 (The section importance is 0.9743), while it is section 20 in AM peak (The section

Table 4 The results of section importance and the ranking in importance when $\alpha = 0.25$

	0:00–24:00		AM peak		PM peak	
Section	I	Ranking	I	Ranking	I	Ranking
1	0.8867	2	0.8114	1	0.8276	2
2	0.9914	1	0.7535	3	0.8833	1
3	0.6147	8	0.4734	15	0.3923	16
4	0.7742	3	0.6019	9	0.5277	7
5	0	26	0.2224	25	0.1371	26
6	0.3197	22	0.4175	21	0.2332	24
7	0.3916	19	0.4583	16	0.3105	19
8	0.2236	23	0.2456	24	0.2646	21
9	0.5764	9	0.4873	14	0.6537	4
10	0.5242	11	0.5948	10	0.4904	9
11	0.4323	18	0.4355	18	0.3479	18
12	0.4575	17	0.4004	22	0.3962	15
13	0.6179	7	0.6597	5	0.5393	6
14	0.6694	5	0.7304	4	0.5016	8
15	0.3723	20	0.4442	17	0.2484	23
16	0.1977	24	0.3610	23	0.1928	25
17	0.4739	15	0.5414	11	0.3919	17
18	0.1758	25	0.0543	26	0.2530	22
19	0.4950	14	0.6253	7	0.4670	10
20	0.4707	16	0.6102	8	0.4409	12
21	0.3632	21	0.5164	12	0.2959	20
22	0.6764	4	0.7539	2	0.6938	3
23	0.5629	10	0.4999	13	0.5826	5
24	0.5212	13	0.6363	6	0.4468	11
25	0.6437	6	0.4297	19	0.4115	13
26	0.5235	12	0.4207	20	0.4004	14

importance is 0.8701), and section 23 in PM peak (The section importance is 0.7810).

5 Conclusion

Traffic congestion usually occurs in some certain sections. The location and capacity of these sections often have a great impact on the road network traffic, and it can be treated as one of the important reasons for leading to traffic congestion. Therefore,

Identifying the critical sections in the road network is an important and critical work in the research on road traffic congestion. On one hand, it plays an important guiding role in urban traffic planning and network transformation. For example, the capacity of critical section is one of the important reasons for leading to traffic congestion. Transportation departments can carry out reasonable network transformation measures pertinently to enhance the capacity of the road network, according to the results of critical section identification. On the other hand, it can provide effective decision support for traffic management measures and residents' travel plans. Traffic management can accurately grasp traffic changes of the road network through which sections are important factors affecting the network traffic states, so as to advances some preventive measures to alleviate traffic congestion. For example, the application of critical section identification in traffic state evaluation or traffic forecasting. These are of great significance to alleviate traffic congestion, improve traffic management and service quality.

The goal of this work is to find the critical section in the road network. Differentiating from the identification methods reported previously, this work not only considers the spatial structure of the road network in the framework of the identification, but also considers the traffic flow characteristics. On the basis of this, section importance is determined by two factors together in this work. One is traffic flow contribution of the section itself, the other is its influence on traffic state of others. Therefore, the critical section is defined as the section which contributes more to the traffic flow of the road network and has greater influence on other sections in the network. Space-time influence and space-time distribution are introduced in this paper to measure the section importance, and that goal is clearly achieved through the section importance ranking by this effort.

For instance, the method applied to in subset of Beijing road network shows that it performs well in finding the critical sections in the test network. It can also be easily applied to other frequency data or experiment scenarios, because they share the common theories foundation. To be honest, there are still some problems to be improved in this method. For example, how to calculate the impact on the road network after the failure of the critical section; How the section importance will change when joining other section factors in the model. All these features make this approach appealing and with plenty of potential for improving. The next steps of this work are as follows:

(1) Continue to verify the effectiveness of this identification method with more date in a larger scale road network, to make sure the universality of the method.
(2) To improve the identification method by considering multiple factors and data sources, for example, section vulnerability, the impact after section failure and so on.
(3) To apply the identification results to the calculation of section weight in the traffic state evaluation, or to apply the identification to traffic forecasting for the critical section.

Declaration of Conflicting Interests The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgements This work was supported in part by the National Key Research and Development Program of China (Grant No. 2017YFC0803900).

References

1. Hossain P, Sayed NSB (2005) Network performance improvement under stochastic events with long-term effects. Transportation 32:65–85
2. Sanso B, Milot L (1999) Performability of a congested urban transportation network when accident information is available. Transp Sci 33(1):68–79
3. Nicholson A, Du ZP (1997) Degradable transportation systems: an integrated equilibrium model. Transp Res Part B: Methodol 31(3):209–223
4. D'Este GM, Taylor MAP (2001) Modelling network vulnerability at the level of the national strategic transport network. J East Asia Soc Transp Stud 4(2):1–14
5. Tu YF, Yang C, Chen XH (2010) Analysis of road network topology vulnerability and critical links. J Tongji Univ (Nat Sci) 38(3): 364–367, 379
6. Hou LW, Jiang F (2004) Study on the relative importance of links in urban roads network. Syst Eng Theory Methodol Appl 13(5):425–428
7. Huang ZF (2009) Order of link trait based on significance and reliability. J Chongqing Jiaotong Univ (Nat Sci) 28(4):742–744
8. Zhao Y, Li H, Wang F (2014) Identifying key sections in urban road network based on k-shell. Syst Eng 32(5):105–110
9. Scott DM, Novak DC, Aultman-Hall L et al (2006) Network robustness index: a new method for identifying critical links and evaluating the performance of transportation networks. J Transp Geogr 14(3):215–227
10. Taylor MAP, D'Este GM, Srinckerhoff P et al (2003) Concepts of network vulnerability and applications to the identification of critical elements of transport infrastructure. Australasian Transport Research Forum, Wellington, New Zealand
11. Zou ZY, Li L, Gao JZ et al (2011) A network efficiency measure for congested transportation networks. In: The 2011 international conference on computational intelligence and software engineering, Wuhan, pp 843–846
12. Liao JQ (2014) An evaluation algorithm for important links of the transportation network based on connect reliability. Mod Comput 11:1–11
13. Zhong R (2013) Research of key nodes and important sections of the road network. Beijing University of Posts and Telecommunications, Beijing
14. Su F, Dong H, Jia L, Sun X (2017) On urban road traffic state evaluation index system and method. Mod Phys Lett B 31(1):1650428
15. Wang YF, He HL (2007) Spatial data analysis method. Science Press, Beijing, pp 56–57
16. Neuman EJ, Mizruchi MS (2010) Structure and bias in the network autocorrelation model. Soc Netw 32(4):290–300
17. Peeters D, Thomas I (2009) Network autocorrelation. Geogr Anal 41(4):436–443
18. Kamarianakis Y, Prastacos P (2005) Space-time modeling of traffic flow. Comput Geosci 31(2):119–133
19. Montis AD, Caschili S, Chessa A (2011) Time evolution of complex networks: commuting systems in insular Italy. J Geogr Syst 13(1):49–65
20. Wang J, Cheng T, Heydecker B et al (2010) STARIMA for journey time prediction in London. In: The 5th IMA conference on math in transport, UCL, London, 12–14 April 2010
21. Box G, Jenkins G (1976) Time series analysis: forecasting and control. Wiley, San Francisco, USA, pp 256–268

22. Chandra SR, Al-Deek H (2008) Cross-correlation analysis and multivariate prediction of spatial time series of freeway traffic speeds. Transp Res Rec 2061:64–76
23. Wu CF (2015) Research on the evaluation of urban road traffic state based on the combination method. Jiangxi University of Science and Technology, Ganzhou, pp 26–30
24. Petter H (2003) Congestion and centrality in traffic flow on complex networks. Adv Complex Syst 6(2):163–176

Public Traffic Passenger Flow Prediction Model for Short-Term Large Scale Activities Based on Wavelet Analysis

Yunqi Jing, Jiancheng Weng, Zheng Zhang, Jingjing Wang and Huimin Qian

Abstract The short-term large scale activities refer to various large-scale activities with a duration of several hours, with features of high peak passenger flow and short gathering time. The analysis of public transport passenger flow characteristics and travel demand prediction for large-scale activities can provide a targeted organization plan for public transportation security in the context of large-scale activities. Based on the smart card data of Beijing, the paper analyzes the spatial-temporal characteristics of passenger flow under the background of large-scale activities. The Discrete-Fourier transform is used to study the frequency domain characteristics of large-scale active passenger flow sequences. Then, through the steps of sampling, decomposition and reconstruction of passenger flow sequence features, the public traffic passenger flow prediction model for short-term large scale activities based on Wavelet analysis was established. And reconstruction steps to establish a short-term large-scale public transport passenger flow forecasting method based on wavelet analysis. The method overcomes the weaknesses that data detail information are ignored in large-scale forecasting during modeling, and improves the stability of forecasting results in short-term forecasting. A case study of Beijing was conducted to validate, and the result shows that the mean absolute percentage error (MAPE) and mean absolute error (MAE) are 0.22% and 1.47%, respectively.

Y. Jing (✉) · J. Weng
Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, China
e-mail: jyqifighting@163.com

Z. Zhang
Beijing Municipal Institute of City Planning & Design, Beijing 100045, China

J. Wang
Beijing Municipal Transportation Operations Coordination Center, Beijing 100161, China

Beijing Key Laboratory of Integrated Traffic Operation Monitoring and Service, Beijing 100161, China

H. Qian
Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming 650504, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_97

Keywords Short-term · Large-scale activities · Wavelet analysis · Small-scale time domain · Passenger flow forecasting

1 Introduction

However, large-scale activities have the characteristics of short-term traffic and strong convergence and dispersion. The arrival and departure of a large number of passenger flows aggravates the degree of urban road congestion, and public transportation, as the main mode of transportation for short-term large-scale activities, greatly alleviates the traffic pressure around the host cities of large-scale activities. Therefore, it is of great practical significance to scientifically analyze and study the passenger flow arrival and departure characteristics of public transportation during large-scale activities and accurately predict the passenger flow scale of short-term large-scale activities.

The passenger flow of public transportation is greatly affected by random factors, and the passenger flow formed is a complex nonlinear system. Because many factors, including weather, holidays, traffic control measures and location factors, will affect the passenger flow, it is difficult to accurately predict. At present, many researchers at home and abroad have done a lot of research on the mid and long-term passenger flow forecast of public transportation. Wu et al. [1] Taking the distribution of bus passenger flow in the evening of Beijing Olympic basketball stadium as an example, this paper allocates the passenger flow of the line and compares it with the actual passenger flow, and the minimum error reaches 3%. Li et al. [2] analyzed the relevant characteristics of passenger flow of large-scale activities. The contents and methods of safety monitoring for large-scale activities are discussed, and the key parts of safety monitoring are determined. Yang et al. [3] established a gray model for passenger flow data with the gray forecast algorithm based on the historical OD passenger flow data of all kinds of large-scale activities, and then established a markov correction model to modify the forecast results. Li et al. [4] determines the main factors affecting the distribution of short-term passenger flow by analyzing the features of passenger flow with the data of Beijing metro stations; and then a forecast model has been established by using GD-FNN to predict the short-term entrance passenger Flow.

At present, growing number of researchers study on short-term passenger flow prediction of public transportation. Generally speaking, they mainly focus on the improvement and applicability of forecast methods [5, 6], and there is still a lack of analysis and forecast on the passenger flow scale of short-term large-scale activities. Zhang [7] presented a Kalman filter method developed to forecast short-term passenger flow based on the characteristic analysis, and the solving algorism. Zhou et al. [8] used network modeling technology to predict 4 indicators, including the passenger flow in and out of the station, the passenger flow in the line interval, the transfer passenger flow, the traffic network passenger flow in no more than 15 min as the transient distributed state of the passenger flow, and carried out the pseudo-real pre-test in a short time. Duan et al. [9] taken typical rail transit stations in Beijing as a case study, an ARIMA-GARCH model is established to simulate the forecast

interval (PI), and fit the stochastic volatility of short-term passenger flow. The effect of "sharp peak and heavy tail" is analyzed by using t distribution. The asymmetry volatility effects are addressed by using T-GARCH and E-GARCH models.

Short-term large-scale activity passenger flow has high peak value, short agglomeration time, and is more affected by random factors, with strong temporal variability and uncertainty, which is significantly different from the medium-long term passenger flow forecast and short-term passenger flow forecast that have been studied more in current literature. In order to achieve the accuracy, real-time and reliability of short-term large event passenger flow forecast, this study established the short-term large event public transport passenger flow forecast method based on wavelet analysis, based on the card transaction data of Beijing public transport and combining with the rule of short-term large event passenger flow arrival and departure. This method overcomes the shortcoming of ignoring the data details in the modeling of large-scale forecast and improves the stability of the forecast results on the minute-scale. This study provides a new and reliable method for passenger flow forecast of short-term large-scale activities in megacities.

2 Spatial-Temporal Distribution of Large Scale Passenger Flow Based on Smart Card Data

As to the short-term large-scale activities, the duration is relatively short, with the characteristics of high peak arrival and departure passenger flow, short gathering time, fixed starting and ending time of the activities and the activity place. Therefore, the audience arriving at the scene before and after the beginning of the activities has a similar entry rule. The calculation method of passenger flow demand in the area affected by traffic during large-scale activities can be described as follows:

$$D = S + O + T \tag{1}$$

where, D is the total traffic demand during large-scale activities (passenger flow of the region during large-scale activities); S is traffic demand of social background; O is the induced traffic demand for large-scale activities (traffic demand caused by large-scale activities); T is the travel volume transferred by traffic demand management measures during special large-scale activities.

In order to better explore the spatial and temporal influence of passenger flow arrival and departure on urban traffic system under the background of large-scale activities, smart card data of certain bus lines in Beijing was selected for characteristic analysis. After preliminary data screening, Smart card data is extracted from a concert held in *Wukesong arena* in April 2016, and ten stations of bus line one, in Beijing. Ten bus stops from *Laoshan* to *Cuiwei* south station were selected as research objects to analyze the passenger flow characteristics during the large-scale activities.

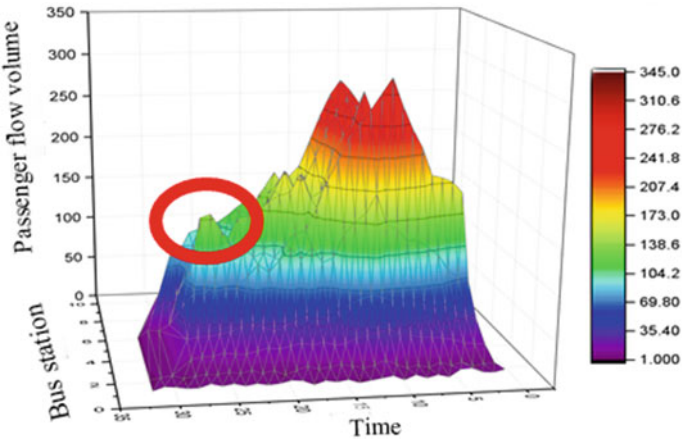


Fig. 1 Spatial-temporal distribution of large scale passenger flow based on smart card data

The passenger flow of activity participants is superimposed with the commuter flow, which is a typical pattern of the passenger flow of large-scale activities. Figure 1 shows a three-dimensional coordinate system with x-axis as the temporal dimension, marking 35 observation groups from the beginning of the activity to the end of the activity; The y-axis is the spatial dimension, and each coordinate represents a station, where "0" represents the starting station of *Laoshan* bus station, "10" represents the south station of *Cuiwei* intersection, the 10th stop of bus line one, and "4" represents the west station of *Shagou* intersection nearest to *Wukesong arena*. The z-axis is passenger flow.

Figure 1 shows that with the increase of y value, z value also increases correspondingly, indicating that the passenger flow of the station located in the center also decreases correspondingly. In terms of time dimension, the first passenger flow peak appears near (10, y, z). During this period, the passenger flow of commuter is superimposed with the passenger flow of large-scale activities, so as to obtain the peak passenger flow. Especially in the vicinity of station 4, the passenger flow gets the maximum. Considering the superposition, it is difficult to analyze the characteristics of passenger flow.

Another peak of passenger flow appears near (26, y, z), as shown in the red circle marked part in Fig. 1. This peak is significantly higher than the passenger flow of the nearby stations during this period, and the passenger flow is mainly dominated by the passengers from large activities.

In order to better explore the spatial-temporal distribution of passenger flow during large-scale activities, the average volume of the passenger flow in recent weeks is separated based on the Formula (1). Thus, the diagram of the Spatial-temporal distribution of passenger flow triggered by large scale activities is obtained, as shown in Fig. 2.

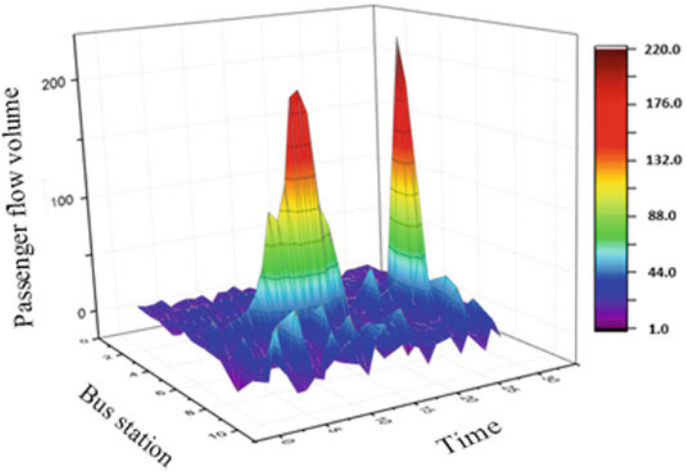


Fig. 2 Spatial and temporal distribution of passenger flow triggered by large scale activities

It can be seen from Fig. 2 that, except for the peak of passenger flow near (10, 4, z) and (26, 4, z), the fluctuation of passenger flow at all other stations was at a low level, indicating that except for station 4, the passenger flow was dominated by daily travel and commuting under the background of large-scale activities, and its landing volume was less affected. The peak of passenger flow appears at station 4 nearest to Wukesong arena. It can be judged that the passenger flow was mainly triggered by large scale activities. In addition, the shape of the wave peak also presents obvious characteristics. when the passenger flow arrives, the wave peak is relatively short and strong, while when the passenger flow leaves, the wave peak is long and thin. It is indicated that the time distribution of the audience arriving at the scene is scattered, while the time of leaving is more concentrated and the peak is larger.

3 Public Transport Passenger Flow Prediction During Short-Term Large-Scale Activities

In the context of large-scale activities, there are many random factors affecting the arrival of passenger flow. The research in the field of forecasting of medium or long time period passenger flow for large-scale activities, such as year, month and day, has been relatively mature. But when these methods are applied to short-term passenger flow sequence forecasting with minute as the observation scale, the error is relatively large, because the detail information of data is ignored in the modeling of large-scale forecasting. The smaller the scale, the more uncertain the sequence and the worse the linear rule. Therefore, in the forecasting of small-scale passenger

flow data, the detailed information in the sequence should be properly processed to ensure the forecasting accuracy.

3.1 Short-Term Passenger Flow Sequence Analysis

3.1.1 Frequency Domain Characteristic Analysis Method

Frequency domain analysis method is widely used in the study of stochastic processes. The traffic volume data collected by direct investigation is time series, and Discrete Fourier Transform (DFT) is adopted to Transform time series into time domain, so as to analyze them in frequency domain.

With the support of Shannon's Sampling Theorem, the numerical implementation of Fourier is feasible. DFT becomes a form of Fourier transform realized by computer. The purpose of DFT transform is to decompose discrete passenger flow sequence signals into orthogonal functions. For any function of period T, it can be expressed as a Fourier series:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos n\omega t + b_n \sin n\omega t) \tag{2}$$

Since the periodicity of the time series of short-time passenger flow data is not obvious, it is assumed that its repetition period region is infinite. On this basis, the spectral density function of $f(t)$ is introduced:

$$F(j\omega) = \int_{-\infty}^{+\infty} f(t)e^{-j\omega t}dt \tag{3}$$

On this basis, its DFT transformation can be obtained:

$$f(t) = \lim_{T\to\infty} \sum_{-\infty}^{+\infty} \frac{F_n}{\omega} e^{jn\omega t}\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(j\omega)e^{j\omega t}d\omega \tag{4}$$

The short-term passenger flow sequence signal can be expressed as a series of sine function integrals of different frequencies through the above DFT transformation, and the frequency domain characteristics of passenger flow sequence signal can be understood through its signal spectrum diagram.

3.1.2 Chaos Characteristics Analysis Methods

Generally speaking, short-time passenger flow signals include the main trend signal and interference signal. Frequency domain analysis shows that there is a random

component in the short-time passenger flow sequence, that is noise, or the inherent deterministic randomness in the short-time passenger flow sequence, that is, chaotic characteristics. Since the chaotic characteristics have a great influence on the forecast method and scale of short-time passenger flow sequence, it is necessary to analyze whether the sequence has inherent deterministic randomness.

In this study, Lyapunov index discriminant method is used to identify chaos characteristics by quantitatively describing the initial value sensitivity of the system, which has a relatively mature theoretical basis. When the Lyapunov index is greater than zero, it indicates that the system has chaotic characteristics. In this paper, the Wolf method based on the theoretical basis of phase space reconstruction of time series was used to calculate the Lyapunov index. The steps are as follows:

3.2 Wavelet Analysis Based Short-Time Passenger Flow Prediction

By analyzing the short-term passenger flow sequence based on the above method, the complex and chaotic components in the short-term passenger flow sequence signal can be stripped out, and the non-stationarity is strong. For short-term large-scale activities, due to the small forecast time scale and numerous influencing factors, it is considered to establish a wavelet forecast method composed of short-term passenger flow sequence decomposition and reconstruction, single signal sequence forecast and simple signal result synthesis.

3.2.1 Schematic of Wavelet Forecasting Method

The principle of short-time passenger flow forecasting method based on wavelet analysis is shown in Fig. 3. Through multi-scale wavelet analysis and fast reconstruction algorithm of Mallat, N step decomposition of the target passenger flow sequence $V = A_0 f(t)$ is carried out to obtain the low-frequency signal $A_N f(t)$ representing the trend of the main passenger flow and the high-frequency signal $D_j f(t)$. The degree N of (j $= 1, 2, \ldots,$ N) deconstruction is determined by the sample variance of the decomposed signals. With the increase of deconstruction times, the low frequency signal of the main passenger flow trend becomes more stable. Therefore, the short-term passenger flow sequence can be decomposed into (N + 1) relatively stable signals, and the ARMA model is used to predict each branch signal separately. Finally, the predicted results of all branch signals are added and reconstructed to obtain the predicted results of the original target passenger flow sequence.

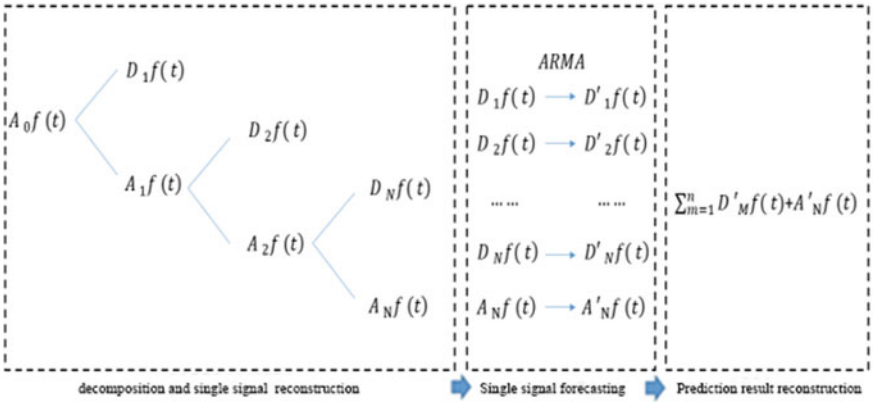


Fig. 3 Schematic diagram of wavelet forecasting method

3.2.2 Method of Signal Decomposition and Reconstruction

The short-time passenger flow sequence is decomposed and reconstructed by Mallat algorithm. The decomposition and reconstruction of Mallat algorithm is the core idea: if $f(t) \in L^2(R)$ known signal resolution in 2^{-j} under the condition of offline approximation $A_j f(t)$, $f(t)$ in resolution 2^{-j+1} of the discrete approximation $A_{j+1} f(t)$ can be used to discrete low-pass filter the $A_j f(t)$ filter.

To $\varphi_{-j,k}(t)$ and $\Psi_{-j,k}(t)$ are $f(t)$ in resolution 2^{-j} under the approximation of the scale function and wavelet function, under the condition of the discrete approximation $A_j f(t)$ and the detail part $D_j f(t)$ can be represented as:

$$\begin{aligned} A_j f(t) &= \sum_{k=-\infty}^{\infty} C_{j,k} \varphi_{-j,k}(t) \\ D_j f(t) &= \sum_{k=-\infty}^{\infty} D_{j,k} \Psi_{-j,k}(t) \end{aligned} \tag{5}$$

In the above formula, $j > 0$. $C_{j,k}$ and $D_{j,k}$ are respectively the resolution 2^{-j} under the trend of low-frequency signal components decomposition coefficient and high frequency detail signal decomposition coefficient. According to tower decomposition idea of Mallat algorithm, $A_j f(t)$ can be further decomposed into the main signal $A_{j+1} f(t)$ and noise signal $D_{j+1} f(t)$, that is,

$$A_j f(t) = A_{j+1} f(t) + D_{j+1} f(t) \tag{6}$$

According to the above two formulas, the decomposition iterative formula of $C_{j,k}$ and $D_{j,k}$ can be obtained, namely, the signal decomposition algorithm:

$$C_{j,k} = \sum_{k=-\infty}^{\infty} h(k-2m)C_{j-1,k}$$
$$D_{j,k} = \sum_{k=-\infty}^{\infty} g(k-2m)D_{j-1,k} \tag{7}$$

where,

$$h(k-2m) = \langle \varphi_{-j,k}, \varphi_{-j-1,m} \rangle$$
$$g(k-2m) = \langle \Psi_{-j,k}, \Psi_{-j-1,m} \rangle \tag{8}$$

On this basis, the synthesis iterative formula of $C_{j,k}$ and $D_{j,k}$, namely the signal reconstruction algorithm, is obtained:

$$C_{j,k} = \sum_{k=-\infty}^{\infty} h(k-2m)C_{j+1,k} + \sum_{k=-\infty}^{\infty} g(k-2m)D_{j+1,k} \tag{9}$$

3.2.3 Single Signal Forecasting Method

After decomposition, the main signal of passenger flow sequence has good stationarity, which can be predicted by the more mature ARMA model. The forecasting model is expressed as:

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \cdots \Phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots \theta_q \varepsilon_{t-q} \tag{10}$$

where, p and q are the order of the forecasting model, determined by the autocorrelation coefficient of the time series sample, and other parameters in the model are obtained by the least square principle. The forecasting result of the original signal is made up of the superposition of each signal.

4 Wavelet Analysis Based Short-Time Passenger Flow Prediction—A Case Study of Beijing

Based on the analysis results of the spatial-temporal characteristics of passenger flow, it can be seen that in the context of large-scale events, the public transport station nearest to the large-scale activities venue has the largest change in passenger flow, while other stations have a small impact. In this paper, the passenger flow data of *Wukesong* subway station near *Cadillac Center (Wukesong arena)* during a concert in July 2018 is selected as the case for analysis, as shown in Fig. 4.

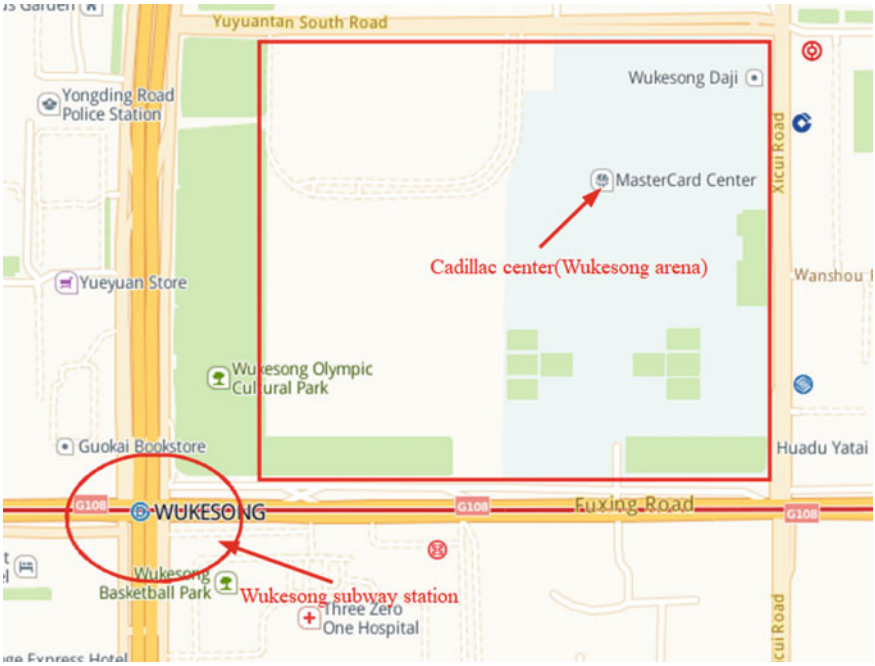


Fig. 4 The location map of Wukesong subway station and Cadillac center(Wukesong arena)

Before and after the start and end of large-scale activities, observers were organized to record the passenger flow that arrived before the beginning and that entered the station after the end of the activity at an observation interval of 5 min to obtain 50 sets of observation data. Taking this as a sample, the forecasting method described above is adopted to analyze and forecast the passenger flow. The observed data are shown in Fig. 5.

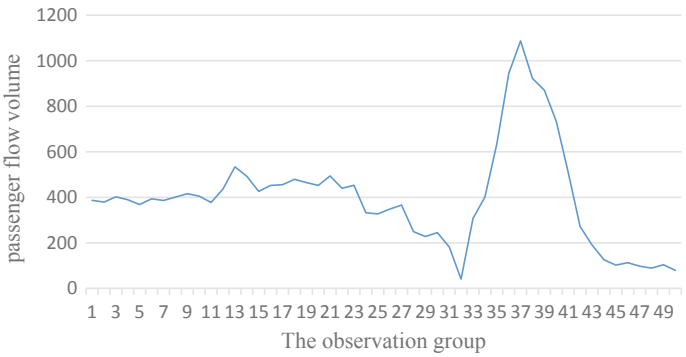


Fig. 5 Actual observation data

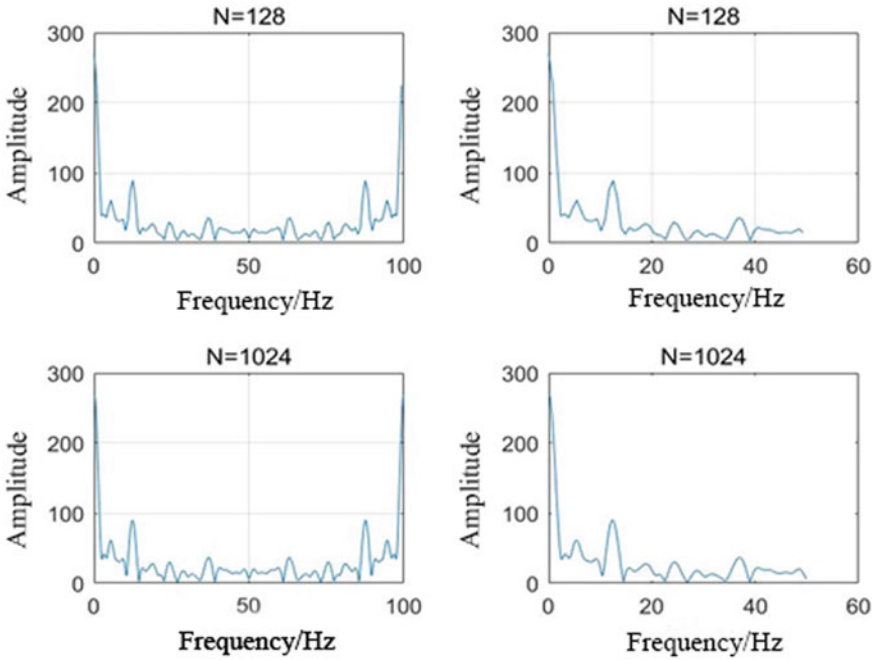


Fig. 6 Signal spectrogram of short-time passenger flow sequence

4.1 Analysis of Passenger Flow Sequence in Frequency Domain

The frequency spectrum of the passenger flow sequence can be obtained by DFT transformation. As shown in Fig. 6, within the frequency range of 0–20 Hz, the signals with the maximum amplitude of the passenger flow sequence are concentrated to collect the passenger flow information of the main public transportation, and the remaining large number of random signal components with small amplitude are distributed in a wide high-frequency domain. Therefore, there are multiple signal components in passenger flow sequence, including complex random signals, and the influence of the above components on passenger flow characteristics is also different, which requires further verification of chaotic characteristics.

4.2 Lyapunov Indexes Calculation and Chaos Identification

Lyapunov, also known as Lyapunov characteristic index, is one of the characteristics used to identify several numerical values of chaotic motion. The time delay index $\tau = 1$ was taken, and the trial algorithm was adopted to calculate the Lyapunov

Tab. 1 Calculation results of Lyapunov indexes

n	Lyapunov indexes
5	0.2435
6	0.2056
7	1.7564
8	1.7542
9	1.7524

exponent when the embedding dimension $n = 3, 4, 5, \ldots, N$ successively until the exponential convergence. The results are shown in Table 1.

According to the Table 1, when the embedded dimension n increases, the Lyapunov index decreases, and when $n > 7$, the Lyapunov index tends to be stable. Therefore, $n = 7$ is taken as the quantitative index to judge the chaotic characteristics of the sequence. On this basis, it can be concluded that among the short-term passenger flow sequence signals, high-frequency signals have an impact on the development trend of passenger flow and cannot be discarded as noise signals. That means that when carrying out passenger flow forecast, its objective impact should be considered.

4.3 Flow Sequence Forecasting

The wavelet forecast method is used to predict the short-term passenger flow sequence. Firstly, db3 function is used as wavelet function and scale function to decompose the original sequence. When the low frequency signal is stable, the original signal is left as a branch signal. The original sequence signal and each branch signal are shown in Fig. 7.

As can be seen from Fig. 7, the stability of the decomposed branch signal is improved greatly compared with the original signal, which is suitable for the prediction by using wavelet analysis. The first 40 groups were used for parameter calibration, and the last 10 groups were used for prediction test, as shown in Fig. 8. And mean absolute percentage error (MAPE) and mean absolute error (MAE) are used to evaluate the prediction results quantitatively.

The verification result shows that MAPE and MAE are 0.22% and 1.47%, respectively. It is indicates that the method of obtaining the branch signals of the original passenger flow sequence by DFT, predicting each branch sequence, and finally superimposing the predicted branch signals has a good effect on improving the accuracy of short-term passenger flow sequence prediction.

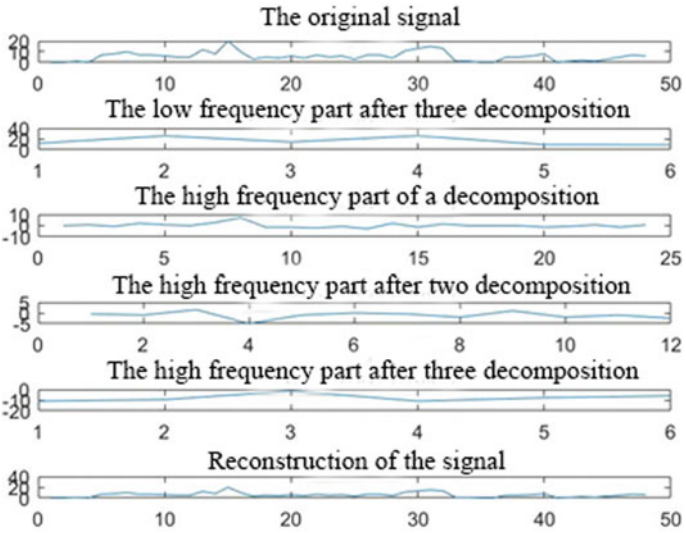


Fig. 7 Original signal and branch signals

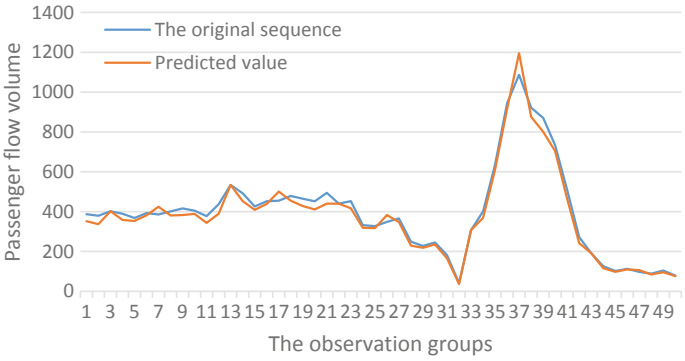


Fig. 8 Prediction results of wavelet method

5 Conclusions

This study analyzes the spatial-temporal characteristics of short-term large-scale activities based on the card swiping data of public transportation. It is shows that the passenger flow induced by large-scale activities is mainly concentrated in the public transportation station nearest to the host place, and the arrival time distribution of passenger flow is scattered, while the departure time is more concentrated and the peak value is larger. By using Discrete Fourier transform, the frequency domain characteristics of original passenger flow data during short-term large-scale activities are extracted, and the prediction method based on wavelet analysis with the process of

decomposition and reconstruction of short-term passenger flow sequence, prediction of single signal sequence, and result synthesis of simple signal is established. A case study of Beijing was conducted to validate, and the result shows that the mean absolute percentage error (MAPE) and mean absolute error (MAE) of this method are 0.22% and 1.47%, respectively, which obviously improves the accuracy and stability of short-term passenger flow forecasting. Because the low-frequency fluctuation of the original sequence is retained, the short-time passenger flow prediction is more accurate, real-time and reliable.

Acknowledgements This research was sponsored by the "Beijing Nova" Program by the Beijing Municipal Science and Technology Commission (Grant No. Z171100001117100), and the National Natural Science Foundation of China (NFSC) (Grant No. 61420106005).

References

1. Yang J, Yang Z (2013) A grey Markov based on large passenger flow real-time prediction model. J Beijing Jiaotong Univ 37(2):119–123
2. Li T, Jin L, Ma Y, Zhu W (2012) Study on method for monitoring and early-warning of passenger flow in large-scale activities. J Saf Sci Technol 8(4):75–80
3. Wu Y, Wei Z, Jian R, Liu X (2011) Study on transit assignment in special event. J Beijing Univ Technol 4:533–540
4. Li, C, Li H, Jiang X et al (2015) Short-term entrance passenger flow forecast at urban rail transit station based on generalized dynamic fuzzy neural networks. Urban Rapid Rail Transit 4:57–61
5. Wei Y, Chen MC (2012) Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. Transp Res Part C Emerg Technol 21(1):148–162
6. Tsai TH, Lee CK, Wei CH (2009) Neural network based temporal feature models for short-term railway passenger demand forecasting. Expert Syst Appl 36(2):3728–3736
7. Zhang C, Song R, Sun Y (2011) Kalman filter-based short-term passenger flow forecasting on bus stop. J Transp Syst Eng Inf Technol 11(4):154–159
8. Duan J, Ding C, Lu Y et al (2017) A prediction approach of short-term passenger flow of rail transit considering dynamic volatility. J Transp Inf Syst 5:68–75
9. Zhou W, Han B, Li D et al (2015) Research and application of short-term forecast model on passenger flow distribution in urban mass transit. Urban Mass Transit 18(2):24–28

The Impact of Subject Diversity on Taxi Transportation System

Wencan Gao, Hao Yue, Bingjian Yang, Mengyu Zhang and Lucheng Zhao

Abstract In order to analyze the impact of different driver types and passenger types on the taxi transportation system, a simulation method is applied to reproduce the process of driver and passenger matching in the taxi market. According to the driver's decision-making way, drivers were divided into three types: searching for the closest passenger in sight, searching for the most profitable passenger per hour by taxi-hailing app, searching for the closest passenger by taxi-hailing app. According to passenger travel characteristics, passengers are classified into ordinary passengers, congested passengers, short-distance passengers and marginal passengers. Formulate passenger taxi generation, passenger disappearance, driver decision, taxi parade, taxi pick-up passengers on board, taxi to transport passenger rules. Select passenger's waiting time, driver's search time, driver's income, taxi's empty rate, taxi's status, number of specific passengers disappeared, ratio of disappeared passengers' number to total disappeared number as evaluation indexes. Research shows that different types of drivers and passengers will have different impacts on the taxi transportation system. Using taxi-hailing app can help improve the taxi transportation system. Searching for the closest passenger by taxi-hailing app mode can improve passenger's actual load efficiency while ensuring the driver's income. It will not screen and eliminate special passengers and improve the fairness of passenger travel. Therefore, it is recommended drivers search for the closest passenger by taxi-hailing app.

Keywords Taxi transportation system · Traffic simulation · Decision-making method · Passenger characteristics · Search radius

W. Gao · H. Yue (✉) · B. Yang · M. Zhang
Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, School of Transportation, Ministry of Transport, Beijing Jiaotong University, Beijing 100044, China
e-mail: yuehao@bjtu.edu.cn

L. Zhao
Nanjing Institute of City & Transport Planning Co., Ltd, Nangjing 210000, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_98

1 Introduction

Taxi traffic is an important part of urban comprehensive transportation system and plays a very important role. According to whether the driver can get the passenger travel information through the taxi-hailing app, the taxi is divided into cruise taxi and online taxi. In terms of traditional cruise taxi, by studying the characteristics of taxi travel based on road network under the condition of fixed demand, the equilibrium model of taxi operation network is constructed [1–3]. The model of taxi service system is established by considering the congestion of road network and the demand elasticity of passengers [4, 5]. In the aspect of online taxi, the simulation model of online taxi operation is established by analyzing the behavior of both the taxi driver and the taxi driver [6–8]. Due to the non-standard taxi-hailing app market and the diversity of individual passengers, some scholars believe that the driver has a high right to know information about passengers, which leads to the refusal of passengers [9, 10].

At present, relevant researches mainly focus on the relationship between supply and demand, simulation, denial of load and other aspects. Most of the research subjects are taxis, and few papers are analyzed from the perspective of diversity of drivers and passengers. This paper analyzes the impact of different types of drivers and passengers on taxi traffic system from the perspective of driver decision diversity and passenger travel diversity, and puts forward the recommended decision-making method for drivers.

2 The Subject of Taxi Transportation System

The subject of the taxi transportation system is driver and passenger. In the taxi system, drivers provide services and passengers are served. Due to the difference of the main purpose and the individual, the diversity exists in both drivers and passengers.

2.1 Driver Classification

The common types of taxi drivers include those who use taxi-hailing app and those who do not. Drivers who do not use taxi-hailing apps rely on the naked eye to search for the nearest passenger. Drivers using the app can have a wider search area and know where passengers are coming from and where they are going. Due to the different ways of drivers' judgment, some drivers prefer to pick up passengers who bring high profits per unit of time, while others prefer to pick up passengers who are nearest to them. Therefore, according to the driver's decision-making way, drivers were divided into three types: searching for the closest passenger in sight,

searching for the most profitable passenger per hour by taxi-hailing app, searching for the closest passenger by taxi-hailing app.

2.2 Passengers Classification

Due to the different purposes and locations of passengers, the characteristics of passenger travel are diverse. Some passengers gathered in the heart of city's downtown area, some passengers are located in remote urban suburbs, some passengers travel during the rush hour, some passengers set off in a clear area during peak hours. According to passenger travel characteristics, passengers are classified into ordinary passengers, congested passengers, short-distance passengers and marginal passengers.

Ordinary passengers who are easy to take a taxi, are mostly located in the center of the region with smooth traffic. Congested passengers refer to passengers in congestion area with inconvenient traffic, drivers may incur excessive congestion costs when picking up the passengers. Short-distance passenger refers to the passenger whose travel distance is too short and only bring the income of starting price. Marginal passengers who located in boundary area, drivers may incur higher costs of empty driving.

3 Simulation Methodology

3.1 Process

In the taxi transportation system, the driver and passenger will go through selecting, matching, picking up and delivering, as shown in Fig. 1.

When the taxi is cruising empty, the taxi randomly selects the target area, travels in this direction and searches passengers for matching; When the taxi matches the passenger successfully, the taxi moves to the starting point of the passenger. After the taxi moves to its matching passenger position and takes a ride, the taxi moves to the target position of passenger travel; When the taxi moved to the passenger's destination, the passenger got off and the taxi returned to cruise.

3.2 Network and Parameter

Urban road network is formed by a number of roads connected, and a number of traffic communities are distributed in the road network. In order to simplify the road network, a checkerboard network is adopted. The straight line is the urban road,

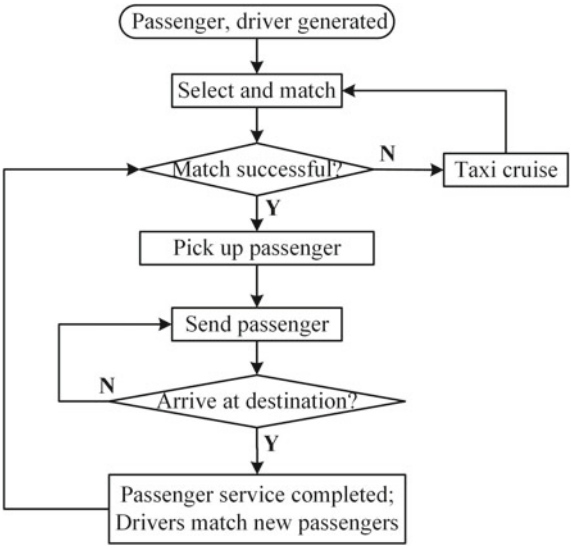


Fig. 1 Running process

namely the taxi driving route, and the intersection point of the straight line is the traffic district, as shown in Fig. 2.

The checkerboard road network consists of 8 * 8 traffic districts, the distance between each district is 1 km, and the region is an 8 km * 8 km road network. Let the unit step 1 s represent the simulation step size, total simulation time is 1 h. At the

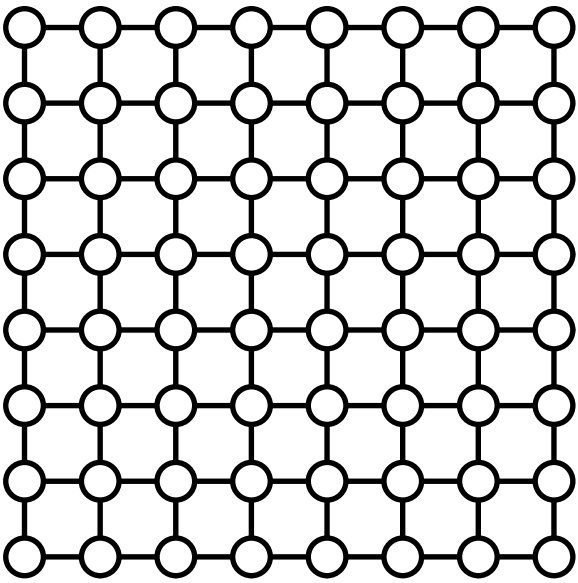


Fig. 2 Simulation road network

beginning, there were 50 passengers, and the region generated 1 person per 1 s on average. The maximum waiting time of passenger is 600 s, and when the maximum waiting time is reached, passengers stop waiting. The number of taxis is always 400.

When drivers don't use taxi-hailing app, passengers can be searched when drivers arrive at the traffic community. When drivers use taxi-hailing app, drivers can search the range by taxi software, search range is 2 km.

When the taxi cruises, the speed of the taxi is 7 m/s. When taxis pick up non-congested passengers, the taxi speed is 10 m/s. When taxis pick up congested passengers, the taxi speed is 5 m/s. Taxi starting price is 14 yuan, unit mileage price is 2 yuan/km, unit wait-time price is 0.5 yuan/min. During the simulation process, various indexes are continuously counted for analysis.

3.3 Evaluation Index

Average waiting time of passengers is an important factor to evaluate passenger travel satisfaction. Drivers' searching time for passengers and average income are important factors for evaluating the service efficiency of drivers. Taxi empty rate and taxi state proportion are important factors to evaluate the effective mileage of taxi. The number of disappearing people in a specific group and the ratio between the number of disappearing people in a specific group and the total number of disappearing people are the evaluation indexes to evaluate the selection of drivers for a specific group. Therefore, the above indicators are selected as evaluation indicators.

Taxi empty rate is the ratio of taxi no-load mileage to total mileage. No-load mileage including cruise distance and taxi driving distance to the departure point for booking passengers.

Average income of drivers refer to the average earnings of a driver in a certain period.

According to the status of drivers and passengers, drivers are divided into cruise status (state = 0), reservation status after successful matching (state = 1), and delivery status after passengers get on the bus (state = 2). Passengers are divided into unreserved waiting state (state = 0), waiting state after successful matching (state = 1), and carrying state after boarding (state = 2).

The vanishing number of specific groups refers to the number of passengers who refuse to continue to wait because they reach the upper limit of waiting time.

4 Results and Discussions

4.1 Analysis of Ordinary Passengers

Taxi drivers using taxi-hailing app can effectively improve the efficiency of the taxi system, reduce the difficulty of ordinary passengers' travel. In terms of passenger's waiting time, drivers who use taxi-hailing app are significantly less than those who do not. In terms of driver's search time, taxi-hailing app can reduce the time for drivers to find passengers. In terms of driver's income, the drivers who use taxi-hailing app have similar earnings, both of which are better than those who do not use taxi-hailing app. In terms of taxi's empty rate, after using taxi software, the empty rate is significantly reduced, and the drivers who search for the closest passenger by taxi-hailing app has the lowest empty rate. As for the reservation state (state = 1), the driver who search for the most profitable passenger has the highest proportion. As for the board state (state = 2), the driver who search for the closest passenger by taxi-hailing app has the highest proportion, as shown in Table 1.

Table 1 Evaluation form for ordinary passengers

Type of drivers	Passenger's wait time (s)	Driver's search time (s)	Driver's income (yuan)	The empty rate of taxi (%)	The proportion of taxi state (state = 0) (%)	The proportion of taxi state (state = 1) (%)	The proportion of taxi state (state = 2) (%)
Drivers who search for the closest passenger in sight	124.3	70.6	120.4	20.3	20.0	0.9	79.1
Drivers who search for most profitable passenger by taxi-hailing app	108.2	6.1	127.9	17.5	4.7	13.2	82.1
Drivers who search for the closest passenger by taxi-hailing app	112.7	6.1	128.1	11.1	4.7	7.7	87.6

4.2 Analysis of Congested Passengers

The influence of congested passengers on the taxi system is analyzed by setting the proportion of different congested passengers and ensuring the other passengers are ordinary passengers. The simulation results are shown in Fig. 3.

In terms of passenger's waiting time, with the increase of the proportion of congested passengers, the vehicle speed decreases, the time needed to carry passenger increases. Therefore, the waiting time increases and finally tends to be stable. The driver who search for the closest passenger in sight is superior to other modes in most cases, as shown in figure (a).

In terms of the search time of drivers, congestion does not affect the search time of drivers, which is determined by the search radius and the driver's own choice. Therefore, the drivers who use taxi-hailing app have basically the same effect, which is better than drivers who don't use, as shown in figure (b).

In terms of driver's income, passengers bear the cost of vehicle congestion after getting on the bus, so the increase of crowded passengers will not reduce taxi driver benefits. With the increase of the proportion of passengers in the congested area, the benefits of drivers in the three modes are all improved, and the benefits of drivers who search for the closest passenger by taxi-hailing app are the best, as shown in figure (c).

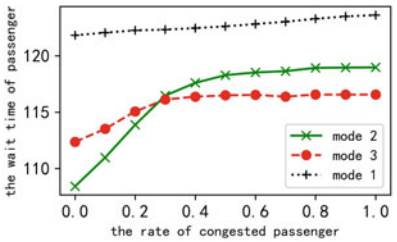
In terms of taxi's empty rate, the increase of crowded passengers makes more vehicles in a slow state of carrying passengers, so the empty rate of the drivers who search for the closest passenger in sight is in a steady state, while the other mode are in a slow state of decline. And the driver who search for the closest passenger by taxi-hailing app has the minimum value, as shown in figure (d).

In terms of the disappeared number of congested passengers, with the increase of congested passengers, the elimination of congested passengers gradually increases. The drivers who search for most profitable passenger deliberately select "high-profit" passengers, thus eliminating the most congested passengers. Drivers who search for the closest passenger by taxi-hailing app have a wide search scope, but limited access to passenger information, limited ability to screen passengers, and the number of passengers eliminated due to congestion is the least, as shown in figure (e).

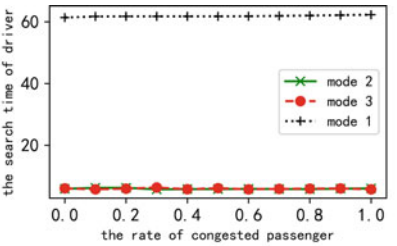
In terms of the number of passengers disappearing due to congestion taking up the total number of passengers disappearing, drivers who search for most profitable passenger deliberately select and display more clearly, as shown in figure (f).

In terms of the proportion of taxi in each state, the proportion of congested passengers has a small impact. The search scope of the unexposed mode is small and more time is spent on cruising. As for the driver who search for most profitable passenger, and the reserved passenger may be far away, thus consuming more time in the reserved state (state = 1). The driver who search for the closest passenger by taxi-hailing app, in the boarding state (state = 2) for the longest, as shown in figure (g).

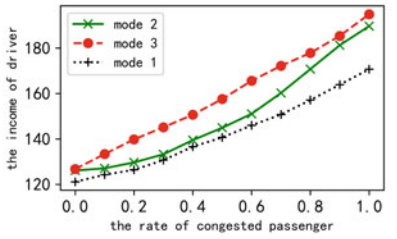
(mode 1 represent the driver who searching for the closest passenger in sight, mode 2 represent the driver who searching for the most profitable passenger per hour by



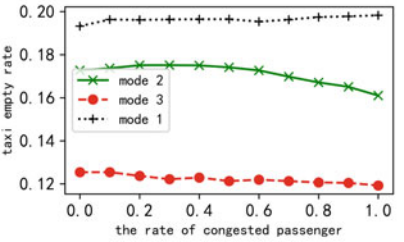
(a) Waiting time of passengers



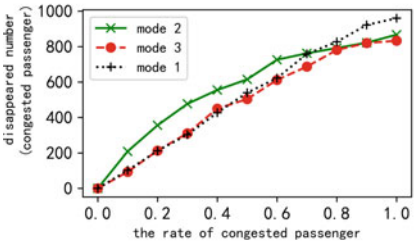
(b) Search time of drivers



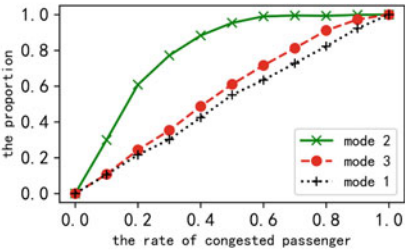
(c) Income of drivers



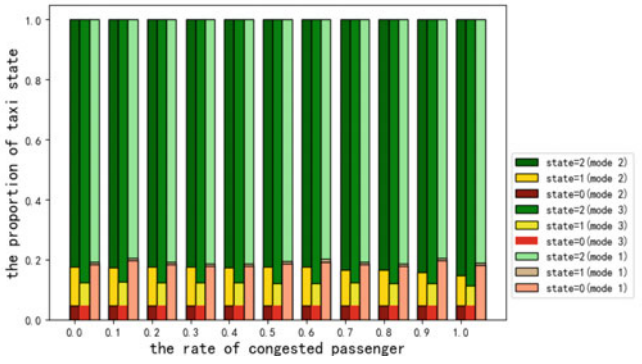
(d) Taxi empty rate



(e) Number of passengers disappearing from Congested areas



(f) The proportion of passengers disappearing from congested areas



(g) The proportion of taxis in each state

Fig. 3 Analysis of congested passengers

taxi-hailing app, mode 3 represent the driver who searching for the closest passenger by taxi-hailing app.)

4.3 Analysis of Short-Distance Passengers

The influence of short-distance passengers on the taxi system is analyzed by setting the proportion of different short-distance passengers and ensuring the other passengers are ordinary passengers. The simulation results are shown in Fig. 4.

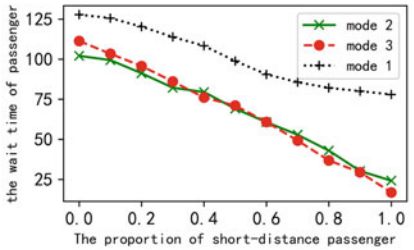
In terms of passenger's waiting time, short-distance passengers travel a short distance and each order takes a short time. With the increase of short-distance passengers, the frequency of taxi transporting passengers increases and the waiting time decreases. In addition, the driver who use taxi-hailing app is better than the other, as shown in figure (a).

In terms of search time of drivers, short distance passengers have little influence on the search time for driver who use taxi-hailing app. With the increase of short-distance passengers, the completion time of each order for drivers for the driver who search closest passenger in sight is shortened, but the search time is longer due to the limitation of search scope, as shown in figure (b).

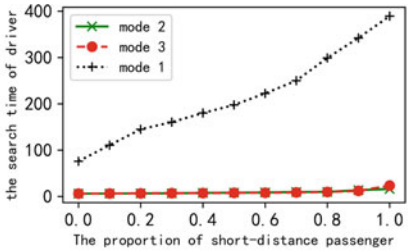
In terms of driver's income, although the order for short-distance passengers only has a starting price, each order can be completed quickly, and the number of orders accepted by drivers in unit time increases greatly. Therefore, with the increase of the proportion of short-distance passengers, the earnings of drivers in the three modes are all improved. The earnings of drivers who use taxi-hailing app higher than the other, as shown in figure (c).

In terms of taxi's empty rate, short-distance passengers have a shorter distance to complete each order, and the taxi will have more mileage to search for new passengers. Therefore, with the increase of short-distance passengers, the taxi empty rate increases gradually. The empty rate of driver who search for the closest passenger by taxi-hailing app is the lowest, followed by the driver who search for most profitable passenger by taxi-hailing app, as shown in figure (d).

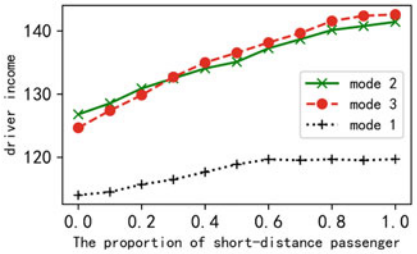
In terms of the number of short-distance passengers disappearing, short-distance passengers bring higher income per unit time, so the driver who search for most profitable passenger will give priority to select such passengers, and fewer short-distance passengers will be eliminated in this mode. And using taxi-hailing app can enable taxis to transport more passengers. As the driver who search passenger in sight, the number of passengers disappeared slowly increased, and the driver who search for the closest passenger by taxi-hailing app was in a stable state at first and then gradually decreased, that is, the supply and demand of taxis changed from balance to oversupply, indicating that a large number of short-distance orders could reduce the number of taxis needed by passengers due to their rapidity, as shown in figure (e).



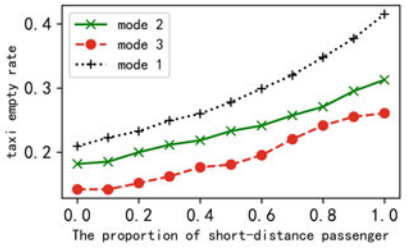
(a) Waiting time of passengers



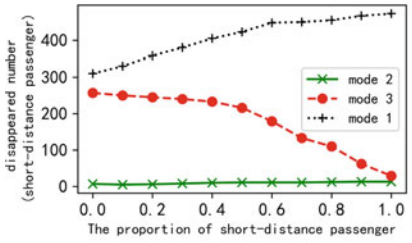
(b) Search time of drivers



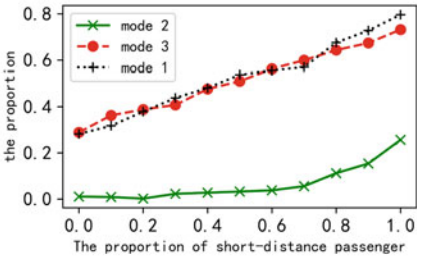
(c) Income of drivers



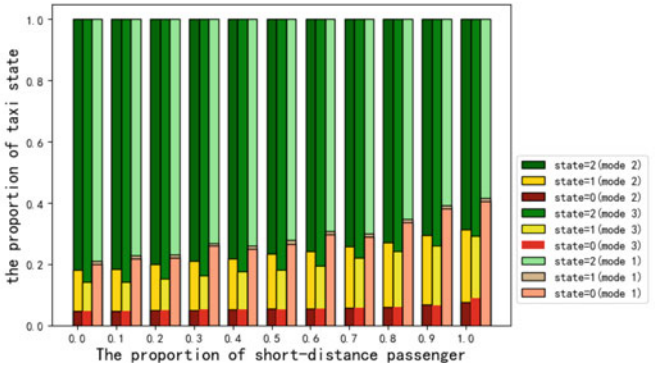
(d) Taxi empty rate



(e) Number of passengers disappearing from Short-distance travel



(f) The proportion of passengers disappearing from short-distance travel



(g) The proportion of taxis in each state

Fig. 4 Analysis of the impact of short-distance travel on passengers

In terms of the proportion of short-distance passengers disappearing in the total number of passengers disappearing, the preference of driver who search most profitable passenger for short-distance passengers is shown more clearly, as shown in figure (f).

In terms of the proportion of taxi states, the influence of short-distance passengers is more obvious. With the increase of short-distance passengers, taxi mileage decreases gradually, the proportion of empty driving increases gradually, and the proportion of passengers decreases. The driver who search for the closest passenger by taxi-hailing app, in the boarding state (state = 2) for the longest, as shown in figure (g).

(mode 1 represent the driver who searching for the closest passenger in sight, mode 2 represent the driver who searching for the most profitable passenger per hour by taxi-hailing app, mode 3 represent the driver who searching for the closest passenger by taxi-hailing app.)

4.4 Analysis of Marginal Passengers

The influence of marginal passengers on the taxi system is analyzed by setting the proportion of different marginal passengers and ensuring the other passengers are ordinary passengers. The simulation results are shown in Fig. 5.

In terms of waiting time of passengers, marginal passengers are more difficult to be found by drivers than ordinary passengers. Therefore, the waiting time increases with the increase of marginal passengers, and the driver who use taxi-hailing app is better than the other, as shown in figure (a).

In terms of search time of drivers, passengers in marginal areas are far away, which makes it more difficult for drivers to search. Therefore, as the proportion of marginal passenger increases, the search time of drivers increases, as shown in figure (b).

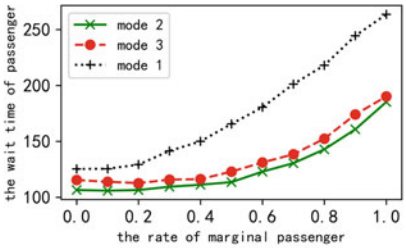
In terms of income of drivers, marginal regions bring a negative impact on the income of drivers, and the more passengers there are, the lower the earnings of drivers will be, as shown in figure (c).

In terms of taxi's empty rate, drivers will waste more mileage searching for marginal passengers, so as the proportion of marginal passenger increases, the taxi empty rate increases gradually, as shown in figure (d).

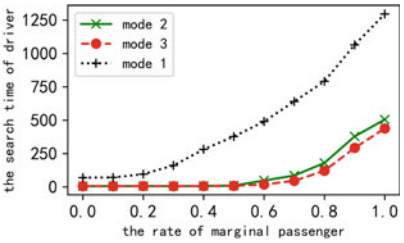
In terms of the number of disappearing passengers in marginal areas, as the number of passengers in marginal areas increases, the number of disappearing such passenger increases, as shown in figure (e).

In terms of the total number of passengers disappearing in marginal areas, the driver who search for the closest passenger by taxi-hailing app is higher than the driver who search for the closest passenger by taxi-hailing app, indicating that drivers will deliberately eliminate passengers in marginal areas, as shown in figure (f).

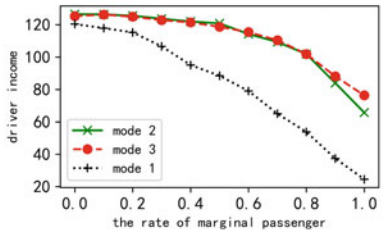
In terms of the proportion of taxi status, passengers in marginal areas have a serious negative impact on the traffic system. The empty driving status of taxis



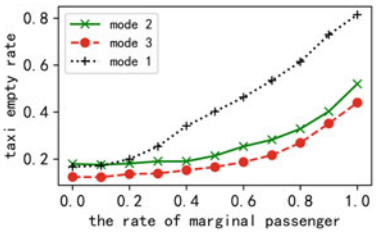
(a) Waiting time of passengers



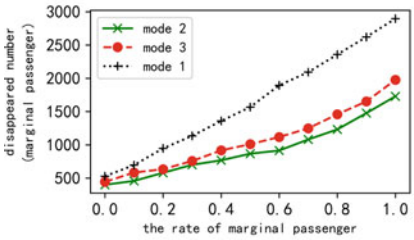
(b) Search time of drivers



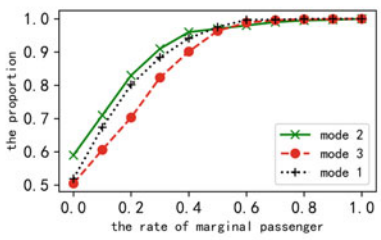
(c) Income of drivers



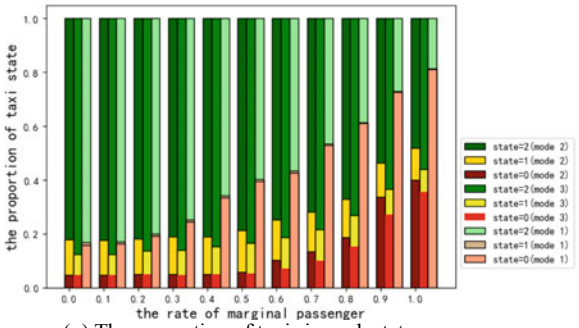
(d) Taxi empty rate



(e) Number of passengers disappearing from Boundary area



(f) The proportion of passengers disappearing from boundary area



(g) The proportion of taxis in each state

Fig. 5 Analysis of the impact of boundary area passengers

increases rapidly with the increase of passengers in marginal areas. Due to the largest randomness of blind cruise in the unpublicized mode, the deterioration effect on the driver who don't use taxi-hailing app is the most serious, as shown in figure (g).

(mode 1 represent the driver who searching for the closest passenger in sight, mode 2 represent the driver who searching for the most profitable passenger per hour by taxi-hailing app, mode 3 represent the driver who searching for the closest passenger by taxi-hailing app.)

5 Conclusions

(1) Using taxi-hailing app can improve the operation efficiency of the taxi system. Compared with other modes, taxi drivers with no use of taxi-hailing app have the longest cruising status, the longest reservation status and the longest passenger carrying status. On the whole, the driver who search for the closest passenger by taxi-hailing app improves the actual load rate of passengers, and the effect is optimal. On the premise of ensuring the income of drivers, the driver reduces the option of passengers, and improves the fairness of passengers' travel.

(2) A driver is a rational person, but there is a limit to rationality. The driver thinks the highest profit per unit time is the highest profit overall, but this is not correct. The driver's "limited rationality" is amplified in the case of information disclosure, and the additional cost caused by empty driving is not fully considered when selecting passengers, so the negative effect is amplified, resulting in the drivers who search for most profitable passenger by taxi-hailing app slightly worse than that of driver who search for the closest passenger by taxi-hailing app.

(3) Different types of passengers have different impacts on the taxi traffic system. Congested passengers reduce the average speed of taxis, the service time of each order increases, the empty driving rate of taxis decreases, and the number of passengers disappears increases. However, since the cost of congestion is borne by passengers, the benefits of drivers increase rather than decrease. Short-distance passengers have shorter journeys, so the service time of each order is short. The growth of cruising time of drivers leads to the increase of taxi empty rate, and the increase of single service increases the income of drivers. And when the short distance passengers too many can be appropriate to reduce the number of taxis to prevent oversupply. The location of passengers in marginal areas is remote, which makes it inconvenient for drivers to search and increases the cost of cruise for drivers. The more passengers there are in marginal areas, the more significant negative effect it will have on the taxi system.

(4) In order to ensure the benefits of drivers and fairness of all kinds of passengers, it is recommended that drivers use taxi-hailing app to find the nearest passengers.

Acknowledge The work is supported by the National Natural Science Foundation of China (Grant No. 71771013, 51338008, 71621001, 71501011), and the Center of Cooperative Innovation for Beijing Metropolitan Transportation.

References

1. Yang H, Wong SC (1998) A network model of urban taxi services. Transp Res Part B 32(4):235–246
2. Yang H, Wong SC, Wong KI (2002) Demand-supply equilibrium of taxi services in a network under competition and regulation. Transp Res Part B 36(9):799–819
3. Yang H, Ye M, Wilson HT (2005) Regulating taxi services in the presence of congestion externality. Transp Res Part A 39(1):17–40
4. Yang H, Wong KI, Wong KI (2001) Modelling urban taxi services in road networks: progress, problem and prospect. J Adv Transp 35(3):237–258
5. Wong KI, Wong SC, Yang H (2001) Modeling urban taxi services in congested road networks with elastic damand. Transp Res B 35(9):819–842
6. Cao Y, Luo X (2016) Equilibrium model of urban taxi service network with the influence of taxi-hailing applications. J Transp Syst Eng Inf Technol 16(02):70–76
7. Cao Y, Li YJ, Luo X (2018) Simulation of taxi passenger travel mode considering the influence of online booking taxi. J Syst Simul 30(02):505–512
8. Du W, Gan HC, Liu BQ (2016) Simulation model for taxi service market equipped with taxi-calling apps. J Transp Syst Eng Inf Technol 16(05):90–96
9. Yuan C, Wu Q, Wei D, Wu D (2014) Optimal modeling and equilibrium mechanism of taxi market with consideration of service refusal. China J Highw Transp 27(6):91–97
10. Yin Y, Yang X (2016) The impact of emerging taxi taking apps on market equilibrium in China. J Dalian Univ Technol 37(2):65–70

Driving Behavior Characteristics on Urban Expressway On- and Off-Ramp by Simulation

Cai Xin, Zhong Yi, Zhao Yong and Mao Yan

Abstract Traffic bottlenecks are easy to form on ramps, which restrict the capacity of urban expressway. The paper studied vehicle driving behavior characteristics on on- and off-ramps by simulator test, the utilization of acceleration section of on-ramps and that of deceleration section of off-ramps, speed transition on ramps, the merging position on on-ramps and the departure position on off-ramps were analyzed. The research would provide theoretical foundations for ramp alignment evaluation, traffic flow control on ramps and coordinated ramp control strategy optimization. It was discovered that vehicle acceleration did not distributed uniformly on ramps. Vehicle driving speed increases before acceleration section on on-ramps and it decreases before deceleration section on off-ramps, which indicates the acceleration and deceleration sections were not used adequately. The analysis of vehicle driving speed on the key positions showed that speed transition on off-ramps was better than that on on-ramps. The merging point where vehicle coming from ramps enter the expressway traffic flow clusters near the end of acceleration section, and the departure point where vehicle coming from expressway enters the ramp clusters on the half part of transition section. This phenomenon indicates that the transition section of on- and off-ramp was not used with expectation.

Keywords Urban expressway · On- and off-ramps · Driving behavior characteristics · Optimize control

C. Xin · Z. Yi · Z. Yong
Shenzhen Traffic Facilities Construction Center, Shenzhen 518040, Guangdong, China
e-mail: 154008640@qq.com

Z. Yi
e-mail: 40582019@qq.com

Z. Yong
e-mail: 306835736@qq.com

M. Yan (✉)
Road Safety Research Center, Research Institute of Highway,
Ministry of Transport, Beijing 10088, China
e-mail: yan.mao@rioh.cn

Key Laboratory of Road Safety Technology, Ministry of Transport, Beijing 10088, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_99

1 Introduction

Urban expressway provides a fast, convenient, comfortable and safe environment for drivers, it improves urban traffic efficiency and alleviates urban traffic pressure. However, traffic bottlenecks are easy to form on ramps, which restrict the expressway traffic capability [1–4].

Many researches had been carried out on ramp traffic flow control and infrastructure construction [5–9]. EVANS studied the traffic flow merging features of on-ramps [10]. Zheng classified ramp control strategies, studied the difference between them and obtained the application scope of each strategy [11]. Zhu studied ramp style, the style setting regulations and its minimum setting distance [12]. Li investigated the pre-set distance of exit signs of urban expressway under different designed speeds. Few researches had been found to investigate vehicle driving behavior on ramps [13].

Therefore, the paper studied vehicle driving behavior characteristics on ramps by simulator tests, the utilization of acceleration section of on-ramps of that of deceleration section of off-ramps, speed transition on ramps and the merging and departure position on ramps were analyzed. The research was of great significance for ramps evaluation, ramps traffic flow control and coordinated ramp control strategy optimization.

2 Methods

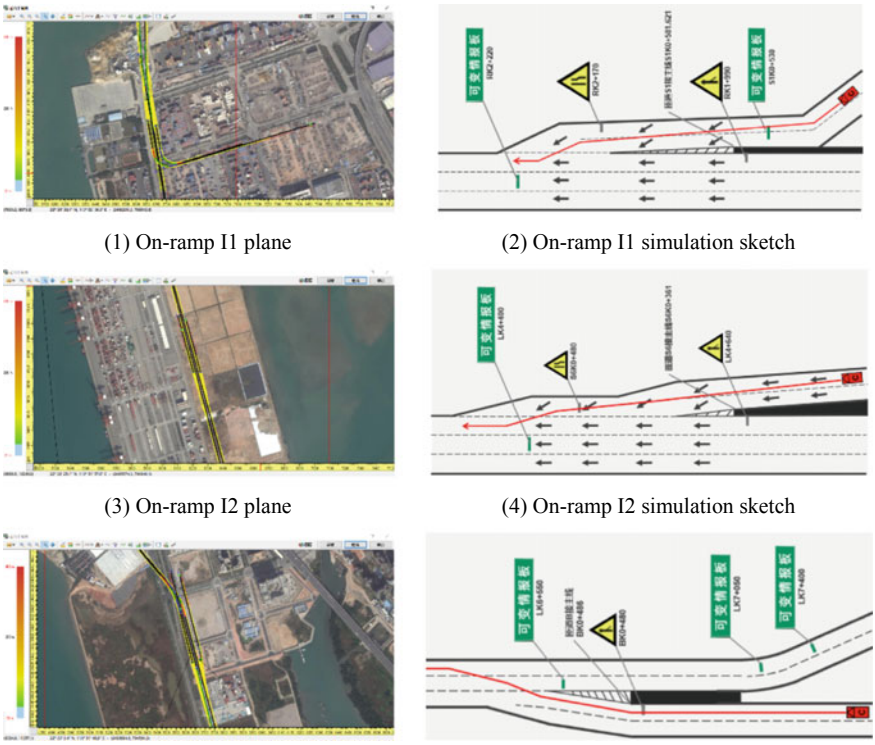
2.1 Simulation Environment

2.1.1 Test Scenario

Three urban expressway on-ramp and three off-ramp sections were established in UC-winRoad environment according to road design information. The on-ramp sections were labelled as I1, I2 and I3 (Fig. 1), and the off-ramp sections were labelled as O1, O2 and O3 (Fig. 2).

2.1.2 Apparatus

The simulation tests were carried out by the research driving simulator with 8-DOFs (Fig. 3). The simulator is composed of motion platform, sphere simulation cabin, projection system, audio system, vehicle simulation system, scene generation system, control platform, power supply system, related auxiliary system and data recording system. The motion platform contains 6-DOFs motion subsystem, X-Table, Yaw-Table and Vibration platform, which could achieve vehicle motion in eight directions. The simulation cabin uses a closed sphere with a diameter of about 4 m and a height of about 3.5 m, in which a simulated vehicle exists and the



(1) On-ramp I1 plane

(2) On-ramp I1 simulation sketch

(3) On-ramp I2 plane

(4) On-ramp I2 simulation sketch

(5) On-ramp I3 plane

(6) On-ramp I3 simulation sketch

Fig. 1 Plane and simulation sketches of urban expressway on-ramps

vehicle could be passenger car or a truck cabin. The simulation cabin is equipped with eight Sim5 W professional projectors, which provide panoramic projection with vertical angle about 40° and horizontal angle about 360°. Vehicle driving simulation is achieved by CarSim or TruckSim. Scene generation and its control is developed by Uc-winRoad. In the Uc-winRoad environment, road models could be constructed in 3D earth, road environment in various conditions could be achieved by changing weather, environment temperature, road adhesion coefficient, slopes, and traffic flow setting. The data recording system transmits and saves hundreds of kinds of data, such as steering wheel and, clutch operation, vehicle driving speed and acceleration, vehicle lane position and so on. Data collected by eye tracker and other physiological equipment could be synchronized recorded.

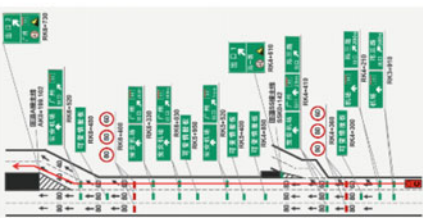
(1) Off- ramp O1 plane

(2) Off-ramp O1 simulation sketch

(3) Off- ramp O2 plane

(4) Off-ramp O2 simulation sketch

(5) Off- ramp O3 plane



(6) Off-ramp O3 simulation sketch

Fig. 2 Plane and simulation sketches of urban expressway off-ramps

Fig. 3 The RIOH researching simulator with 8-DOFs

2.2 Participants and Test Requirement

2.2.1 Participants

18 participants were recruited in this test, whose ages change from 22 to 27 years old and with driving experiences of 1–7 years. All of them had valid China driving license. Their detailed information was listed in Table 1.

2.2.2 Test Requirement

Before the formal driving simulation, the driver was informed of the operation method of the driving simulator, and adaptive driving in the simulator was conducted for each driver. In the formal driving test, the driver was asked to drive along the designed road, which consists of the three on-ramp and three off-ramp sections and urban expressway. The designed driving speed of the expressway is 80 km/h. Drivers were required to drive at a speed not higher than 80 km/h.

2.3 Data Collection and Preprocessing

2.3.1 Data Collection

The research aimed to reveal driving behavior characteristics on on-ramps and off-ramp sections by the utilization analysis of on-ramp acceleration and off-ramp deceleration sections, speed transition analysis of on- and off-ramp, and analysis of the key vehicle merging position on on-ramp and departure position on off-ramp. Therefore, vehicle acceleration, accelerator pedal and brake pedal operation, vehicle driving speed and trajectory on on-ramps and off-ramps were selected and used in the following analysis.

2.3.2 Data Preprocessing

Ramps can be divided into velocity changing section and transition section (Fig. 4) according to Specification for Design of Highway Route (JTG D20-2017). The length of these two parts was regulated according to the designed speed of expressway. There hence, the velocity changing and transition sections of each ramp were picked out in this study.

Table 1 Participants information

No.	Driving license	Sex	Age	Driving experience	No.	Driving license	Sex	Age	Driving experience
SD01	C1	Male	25	3	SD10	C1	Male	25	2
SD02	C1	Male	23	4	SD11	C1	Male	24	2
SD03	C1	Male	26	4	SD12	C1	Male	24	1
SD04	C1	Male	24	3	SD13	C1	Male	23	3
SD05	C1	Male	23	2	SD14	C1	Male	26	4
SD06	C1	Female	25	1	SD15	C1	Male	24	2
SD07	C1	Male	26	5	SD16	C1	Male	24	4
SD08	C1	Male	27	7	SD17	C1	Female	22	1
SD09	C1	Male	26	4	SD18	C1	Female	24	1

(1) Direct single lane (2) Parallel single lane (3) Direct double lane

Fig. 4 Urban expressway ramp style

3 Driving Behavior Characterizes on Urban Expressway On- and Off-Ramp

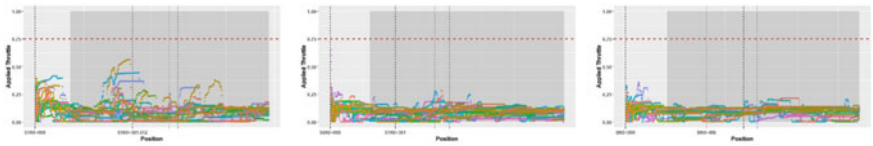
As the link between urban expressway and surface road, ramp plays the role of transition and cushion. While the ramp can be used effectively, vehicles could merge fluently with the traffic flow on expressway or could depart safely from expressway traffic flow. Otherwise, the traffic safety would be influenced.

3.1 The Utilization Analysis of On-Ramp Acceleration and Off-Ramp Deceleration Section

To use the on-ramp acceleration section and off-ramp deceleration section effectively, accelerator pedal and brake pedal were operated to adjust vehicle driving velocity, by which to reduce the gap between vehicle driving speed and road designed speed when vehicle arrives at the expressway.

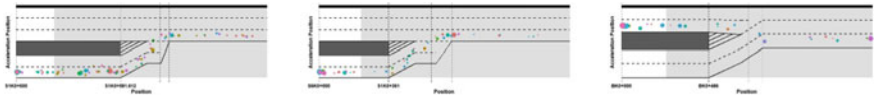
3.1.1 The Utilization Analysis of On-Ramp Acceleration Section

While drives on on-ramps (I1, I2 and I3), drivers' operation on accelerator pedal was shown in Fig. 5. Limited by topography, drivers' operation on accelerator pedal was quite gentle. There is hardly any rapid acceleration, and vehicle speed increase was generally achieved by multiple gentle operation on accelerator pedal. The position where acceleration operation happens was shown in Fig. 6. Acceleration operation discretely distributed on the whole on-ramp, but most of it happens before and on acceleration section.



(1) On-ramp I1 (2) On-ramp I2 (3) On-ramp I3

Fig. 5 Drivers' operation on accelerator pedal on on-ramps

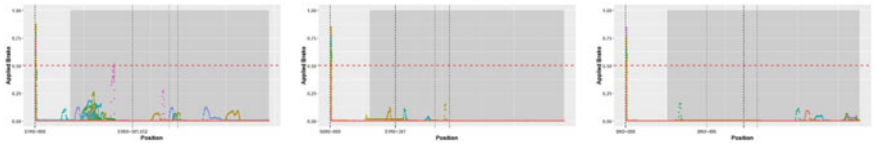


(1) On-ramp I1 (2) On-ramp I2 (3) On-ramp I3

Fig. 6 The acceleration position on on-ramps

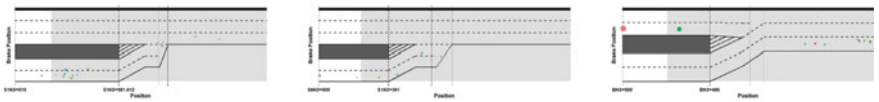
The brake pedal operation on on-ramps was shown in Fig. 7. There was quite few brake operation on on-ramps, and the operation was quite gentle. The position where braking happens was shown in Fig. 8. Few brake operation happens before acceleration section and at the end of transition section, by which vehicle driving speed was adjusted with small amplitude.

Vehicle acceleration on on-ramps were shown in Fig. 9. On the whole on-ramp, vehicle acceleration distributed unevenly. The value of acceleration changes between -3.0 and 3.0 m/s^2. Vehicle acceleration changes greatly before and on acceleration section. Vehicle driving speed increases before acceleration section. There hence, the utilization of on-ramp acceleration section was not satisfied in this research.



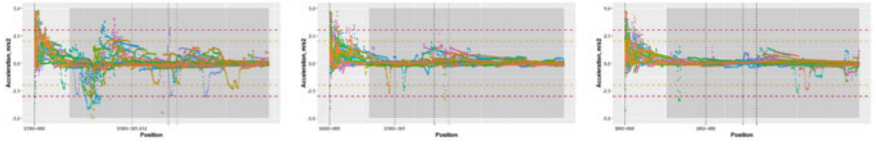
(1) On-ramp I1 (2) On-ramp I2 (3) On-ramp I3

Fig. 7 Drivers' operation on brake pedal on on-ramps



(1) On-ramp I1 (2) On-ramp I2 (3) On-ramp I3

Fig. 8 The braking position on on-ramps



(1) On-ramp I1 (2) On-ramp I2 (3) On-ramp I3

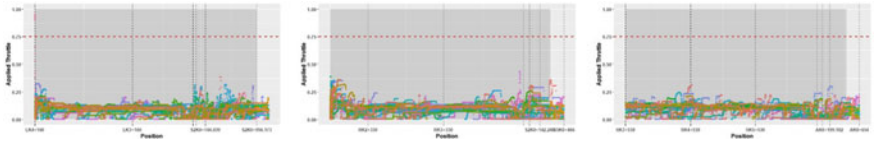
Fig. 9 Vehicle acceleration on on-ramps

3.1.2 The Utilization Analysis of Off-Ramp Deceleration Section

While drives on off-ramps (O1, O2 and O3), drivers' operation on accelerator pedal was shown in Fig. 10. There was gentle acceleration operation on off-ramps. The position where acceleration happens was shown in Fig. 11. The acceleration operation discretely distributed on the whole off-ramps. The operation clusters intensively on deceleration and transition section, which aims to adjust vehicle driving speed with small amplitude.

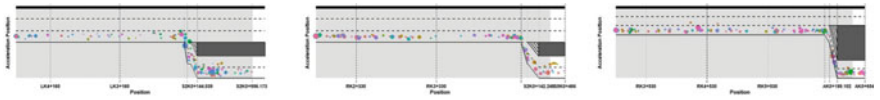
The operation on brake pedal on off-ramps was shown in Fig. 12. Hard braking could be found on transition section and at the end of deceleration section, which accords with the practice. The position where braking happens was shown in Fig. 13. Driver has frequent braking operation on transition section and at the end of deceleration section.

Vehicle acceleration on off-ramps were shown in Fig. 14. On the whole off-ramp, vehicle acceleration does not distributed uniformly. The value of acceleration changes between -3.0 and 3.0 m/s^2. Vehicle driving speed decreases before deceleration section. On deceleration section, vehicle driving speed reduces with large amplitude. There hence, the utilization of off-ramp deceleration section was better than that of on-ramp acceleration section.



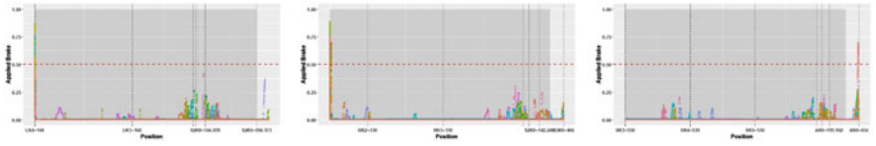
(1) Off-ramp O1 (2) Off-ramp O2 (3) Off-ramp O3

Fig. 10 Drivers' operation on accelerator pedal on off-ramps



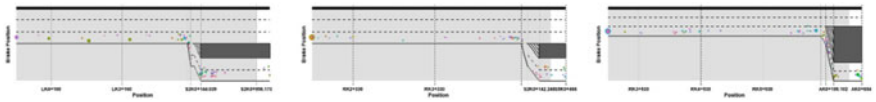
(1) Off-ramp O1 (2) Off-ramp O2 (3) Off-ramp O3

Fig. 11 The acceleration position on off-ramps



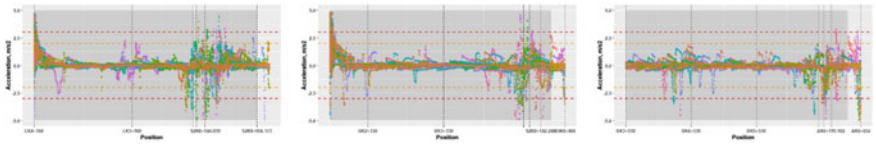
(1) Off-ramp O1 (2) Off-ramp O2 (3) Off-ramp O3

Fig. 12 Drivers' operation on brake pedal on off-ramps



(1) Off-ramp O1 (2) Off-ramp O2 (3) Off-ramp O3

Fig. 13 The braking position on off-ramps



(1) Off-ramp O1 (2) Off-ramp O2 (3) Off-ramp O3

Fig. 14 Vehicle acceleration on off-ramps

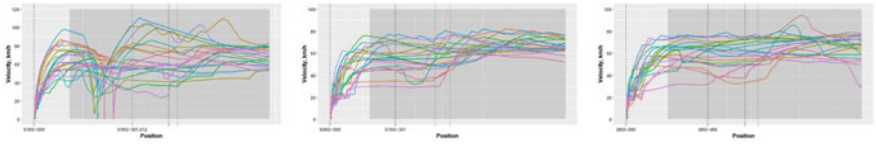
3.2 Speed Transition Analysis of On- and Off-Ramp

Generally, vehicle has low driving speed on surface ground, and the designed speed of expressway is much higher. Therefore, vehicle driving speed increases on on-ramps, and it was very satisfied that vehicle driving velocity at the end of on-ramp is quite close to the expressway designed one. On this situation, vehicle could merge with expressway traffic flow fluently and safely. While drives on off-ramps, vehicle driving speed decreases gradually so as to its driving speed could changes from the higher expressway designed one to the lower surface road designed one. Therefore, the paper utilized vehicle driving speed on ramps to obtain vehicle driving speed on key positions and analyze vehicle speed transition. Vehicle speed transition on ramps was expressed by the error between vehicle actual driving speed and road designed one when vehicle arrived at expressway on on-ramps and surface road on off-ramps:

$$\Delta v = v_{actual} - v_{expect} \tag{1}$$

3.2.1 Speed Transition Analysis of On-Ramp

Vehicle driving speed on on-ramps (I1, I2 and I3) was shown in Fig. 15. The average driving speed and speed on the key positions, the start and end of acceleration section, the median merging position, and the end of transition section were included, were listed in Table 2. Vehicle driving speed on the start of acceleration section was 59.86 km/h, while it at the end of acceleration section (the start of transition section) was 60.12 km/h. Most of vehicles accelerated before acceleration section. On the acceleration section, vehicle driving speed increased a little. Furthermore, vehicles



(1) On-ramp I1 (2) On-ramp I2 (3) On-ramp I3

Fig. 15 Vehicle driving speed on on-ramps

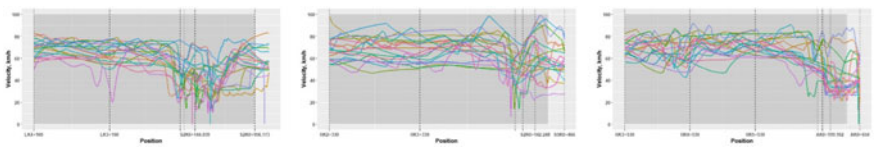
Table 2 The average speed and speed on the key positions of on-ramps

The key position	Vehicle driving speed, km/h			The average speed, km/h	Speed error, km/h
	I1	I2	I3		
The start of acceleration section	65.38	54.96	59.24	59.86	−20.14
The end of acceleration section (the start of transition section)	65.63	54.98	59.74	60.12	−19.88
The median merging position	65.40	53.97	59.72	59.70	−20.3
The end of transition section	64.52	58.81	61.30	61.54	−18.46

tend to accelerated after lane changing. In this study, the designed speed of expressway is 80 km/h, vehicle driving speed on the median merging position was about 59.70 km/h, and it was 61.54 km/h at the end of transition section. Vehicle driving speed while vehicle arrived at the expressway differs a lot from road designed one. The speed transition on on-ramps were not satisfied.

3.2.2 Speed Transition Analysis of Off-Ramp

Vehicle driving speed on off-ramps (O1, O2 and O3) was shown in Fig. 16. The average driving speed and speed on the key positions, the start and end of transition



(1) Off-ramp O1 (2) Off-ramp O2 (3) Off-ramp O3

Fig. 16 Vehicle driving speed on off-ramps

Table 3 The average speed and speed on the key positions of off-ramps

The key position	Vehicle driving speed, km/h			The average speed, km/h	Speed error, km/h
	O1	O2	O3		
The start of transition section	46.15	51.36	56.64	51.38	11.38
The median departure position	42.17	50.18	56.64	49.66	9.66
The end of transition section (the start of deceleration section)	45.74	51.27	55.55	50.85	10.85
The end of deceleration section	46.34	61.89	44.49	50.91	10.91

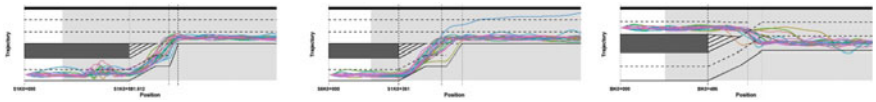
section, the median departure position, and the end of deceleration section were included, were listed in Table 3. Vehicle driving speed on the start of transition section was 51.38 km/h, it on the median departure position was 49.66 km/h, it at the end of transition section was 50.85 km/h, and it at the end of deceleration section was 50.91 km/h. Most of vehicles decelerated before deceleration section. On the deceleration section, vehicle driving speed decreased a little. In this study, the designed speed of surface road is 40 km/h. Speed error on off-ramps was smaller than that on on-ramps. Therefore, speed transition on off-ramps was better than that on on-ramps.

3.3 Analysis of the Key Vehicle Merging and Departure Positions

The merging point is the position where vehicle coming from on-ramps enters the traffic flow in expressway, while the departure point is the position where vehicle coming from expressway enters off-ramps. These two important positions were necessary to predict vehicle driving behavior on expressway or on off-ramps, and the prediction would be helpful to carry out speed control or traffic flow deliver. Therefore, the paper conducted the merging and departure position research.

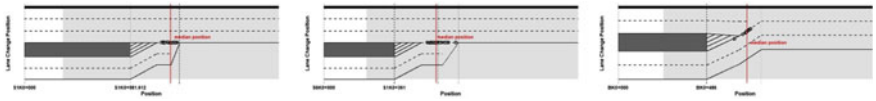
3.3.1 The Key Merging Position from On-Ramp onto Expressway

Vehicle driving trajectory on on-ramps (I1, I2 and I3) was shown in Fig. 17. Extracted vehicle merging position, and the median merging position on each on-ramp was shown in Fig. 18. On I1, the merging position clusters near the end of acceleration section or the start of transition section. Statistical analysis showed that the median merging position was 3 m away from the start of transition section, which indicates



(1) On-ramp I1 (2) On-ramp I2 (3) On-ramp I3

Fig. 17 Vehicle driving trajectory on on-ramps



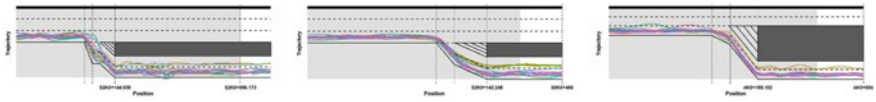
(1) On-ramp I1 (2) On-ramp I2 (3) On-ramp I3

Fig. 18 The merging position on on-ramps

that the transition section was not adequately used. On I2, the merging position also clusters near the end of acceleration section, the distance between the median merging point of the start of transition section was 8 m. The transition section on I2 was not adequately used, too. On I3, vehicles coming from on-ramps merge with traffic flow on expressway at the position near the end of acceleration section. The distance between the median merging position and the start of transition section was 2 m. The transition section on I3 was not used adequately by most of drivers. There hence, in this study, vehicle merging position clusters near the end of acceleration section, drivers did not use the transition section with expectation.

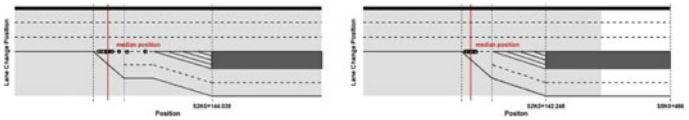
3.3.2 The Key Departure Position from Expressway onto off-Ramp

Vehicle driving trajectory on off-ramps was shown in Fig. 19. Extracted vehicle departure position, and the median departure position on each off-ramp was shown in Fig. 20. On O1, the departure position clusters on the transition section. By statistical analysis, the median departure position on O1 was 24 m away from the start of transition section, and was 26 m away from the end of transition section. On O2, vehicle departures on transition section. The median departure position was 25 m away from the start of transition section and was 55 m away from the end of transition section. There was no need to change lane on O3 while vehicle drives on the right side. The position where vehicle leaves expressway is the start of transition section.



(1) Off-ramp O1 (2) Off-ramp O2 (3) Off-ramp O3

Fig. 19 Vehicle driving trajectory on off-ramps



(1) Off-ramp O1 (2) Off-ramp O2

Fig. 20 The departure position on off-ramps

There hence, in this study, vehicle departure position clusters on the first part of transition section.

4 Conclusions

Traffic flow is quite intensive on on-ramps and off-ramps of expressway, which lead to comprehensive traffic environment. As the traffic bottlenecks, urban expressway traffic capability depends on on-ramps and off-ramps to a large degree. The paper studied vehicle driving behavior characteristics on on-ramps and off-ramps by simulator test, the utilization of acceleration section of on-ramps and that of deceleration section of off-ramps, speed transition on ramps, the merging position on on-ramps and the departure position on off-ramps were analyzed.

It was discovered that vehicle acceleration did not distributed uniformly on ramps. Vehicle driving speed increases before acceleration section on on-ramps and it decreases before deceleration section on off-ramps, which indicates the acceleration and deceleration sections were not used adequately. The analysis of vehicle driving speed on the key positions showed that speed transition on off-ramps was better than that on on-ramps. The merging point where vehicle coming from ramps enter the expressway traffic flow clusters near the end of acceleration section, and the departure point where vehicle coming from expressway enters the ramp clusters on the half part of transition section. This phenomenon indicates that the transition section of on- and off-ramp was not used with expectation. Therefore, it is necessary to set up traffic signs on ramps and optimize traffic flow control on ramps.

Acknowledgements This research was supported in part by the National Key R&D Program of China 2017YFC0804800 (2017YFC0804802), the Traffic Construction Engineering Research Project of the Shenzhen Traffic Facilities Construction Center under Contract MWKH-ky-002, and the Central Public-Interest Scientific Institution Basal Research Fund 2016-9003.

References

1. Li Y, Yao H, Lu Q (2010) Analysis of traffic characteristics of urban expressway ramps. In: Cross-strait intelligent transportation system symposium
2. Li X, Wang Z, Wu B (2015) Study on traffic flow rate characteristics of bottleneck region of urban expressway exit ramp. J Wuhan Univ Technol (Transp Sci Eng) 3:577–581
3. Li T, Li W, Zhou R et al (2001) Traffic characteristics analysis for merging and diverging in acceleration and deceleration lane of expressway. J Highw Transp Res Dev 18(4):89–91
4. Ma J, Zhang L, Li K (2016) Modeling and simulation of traffic flow in urban expressway and ramp intersection. Computer Simulation 12
5. Elefteriadou L, Kondyli A, Washburn S et al (2011) Proactive ramp management under the threat of freeway-flow breakdown. Procedia Soc Behav Sci 16:4–14
6. Rao J (2014) Research on modeling and control method of urban expressway exit ramp area. Zhejiang University
7. Chen X, Tian A, Wang D et al (2007) Research on linkage control strategy of urban expressway entrance and exit ramp. J Traffic Inf Saf 25(6):4–7
8. Zhang H, Cheng W (2012) Design of expressway interchange acceleration and deceleration lane length. J Transp Res 12:130–132
9. Chen Q (2016) Research on safety evaluation of urban rapid route shape design based on driving simulator. China Municipal Eng 3
10. Evans JL, Elefteriadou L, Gautam N (2001) Probability of breakdown at freeway merges using Markov chains. Transp Res Part B (Methodol) 35(3):237–254
11. Zheng J, Dong D, Chen H (2006) Comparative analysis on on-ramp control strategies of urban freeway. J Comput Meas Control 14(2):196–199
12. Zhu S (2004) Discussion on the setting of the entrance and exit of urban expressway. Urban Traffic 2(4):59–63
13. Li A, Li W, Wang W (2006) Research on location setting of exit signs of urban expressway. J Transp Syst Eng Inf 6(5):36–41

Optimization on Design Parameters of Road Longitudinal Slope Based on Truck Dynamics

Cai Xin, Zhong Yi, Zhao Yong and Mao Yan

Abstract In recent years, with major cities turning to the stock development stage, the development and construction of underground space has gradually become an important strategy for sustainable urban development. The construction of underground passages at freight transport hubs will help the separation of passenger and cargo, which will greatly alleviate urban traffic jam and improve transport efficiency. However, due to the restriction on urban land use and construction cost, it is difficult to achieve gentle slope design for all of underground passages, longitudinal slopes with large gradient are unavoidable. Generally, trucks have quite large gross weight, its uphill speed reduction and downhill brake disc temperature rise would affect road capacity and traffic safety greatly. Therefore, it is quite necessary to propose optimal design of longitudinal slopes given the consideration of truck dynamic performance and its driving safety. Truck models were established in TruckSim, they were used to simulate vehicle driving status uphill and downhill then. Speed reduction and brake disc temperature rise were used to obtain suitable slope gradient and length based on simulation. It was concluded that slope gradient and vehicle output power ratio have great influence on truck climbing ability. Using speed reduction of 20 km/h between vehicle uphill original speed and its stable uphill speed for reference, the maximum slope length for uphill with different gradient could be obtained. Slope gradient and vehicle gross weight have giant impact on brake disc temperature rise. Using brake

C. Xin · Z. Yi · Z. Yong
Shenzhen Traffic Facilities Construction Center, Shenzhen 518040, Guangdong, China
e-mail: 154008640@qq.com

Z. Yi
e-mail: 40582019@qq.com

Z. Yong
e-mail: 306835736@qq.com

M. Yan (✉)
Road Safety Research Center, Research Institute of Highway, Ministry of Transport, Beijing 10088, China
e-mail: yan.mao@rioh.cn

Key Laboratory of Road Safety Technology, Ministry of Transport, Beijing 10088, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_100

disc temperature of 200 and 260 °C for reference, suitable slope length for downhill with different gradient were obtained.

Keywords Traffic engineering · Simulation analysis · Vehicle dynamics · Slope gradient · The critical slope length

1 Introduction

In recent years, with major cities turning to the stage of stock development, the development and construction of underground space has become an important strategy for sustainable urban development. The construction of underground passages can rich the transportation system. Furthermore, the construction of underground passages at freight transport hubs will help the separation passenger and cargo, which will greatly alleviate urban traffic jam and improve transport efficiency. However, due to the restriction on urban land use and construction cost, it is difficult to achieve gentle slope design for all underground passages, longitudinal slopes with large gradient are unavoidable. Trucks are widely used in freight transportation, which has quite large vehicle gross weight. While drives uphill, vehicle speed reduction is rather large, the stable uphill driving speed would even be lower than the regulated minimum speed. This phenomenon seriously restricts road capacity and even causes traffic jam. What's worse, bad longitudinal slope design would exceed truck's climbing ability, resulting in collision incidents. While drives downhill, driver needs to brake the vehicle with high frequency so as to control vehicle driving velocity. The frequent operation on brake system would like to cause large temperature rise of brake disc, resulting in heat recession or even brake fade. There hence, to alleviate the impact of longitudinal slope on vehicle driving, the design parameters of longitudinal slope have been studied widely.

Castllio [1] studied vehicle uphill climbing performance under different vehicle laden condition. ASTT [2] researched the uphill climbing ability of trucks with 120 kg/kW output power ratio, then modeled vehicle speed variation along with slope gradient and its length. Zhuang et al. [3] analyzed vehicle speed reduction with respect to slope length, proposed the maximum slope length for slope with different gradient. Dong et al. [4, 5] studied the variation of truck climbing speed on long longitudinal slope section, put forward the design control index of slope gradient and slope length.

In 2004, Western Communication Construction Project [6] proposed quantitative criteria for slope gradient and slope length of continuous longitudinal slope section based on downhill driving safety. Su et al. [7], Han et al. 8] modeled the prediction of truck brake disc temperature, and then put forward the suggestion value of downhill slope gradient, safety slope length and its total length. Wu et al. [9] evaluated the possibility of truck braking failure on downhill with different gradients and lengths using temperature rise model of brake disc. Du et al. [10] established temperature

rise model of brake disc while vehicle drives on long downhill, analyzed the critical gradient, safety risk classification index and evaluation method.

Until now, researches mainly focus on the improvement strategy of uphill or downhill design based on field tests. Test vehicle type and test condition are limited. In practice, many combination of slope gradient and slope length exists, truck type and its laden condition are various. Filed tests can not conducted on all of the test conditions. Therefore, the paper established different kinds of truck models in TruckSim environment, simulate vehicle uphill and downhill driving behavior under different slope gradient, vehicle laden condition and original driving speed. By this method, the paper aims to given a comprehensive consideration of the optimal design of longitudinal slopes.

2 Truck Dynamics and Its Driving Features on Longitudinal Slope

2.1 Force Analysis of Truck

The maximum road longitudinal slope gradient and its length rely on vehicle dynamic performance. Vehicle stable driving speed on uphill can be obtained by force analysis (Fig. 1) and given engine performance parameters.

According to Fig. 1, force balance along vehicle driving direction could be expressed by

$$T = R_w + R_R + R_I \tag{1}$$

where T is vehicle driving force; R_w is air resistance; R_R is road friction resistance; R_I is inertia resistance.

Given the consideration of vehicle output power ratio U, vehicle dynamic factor D can be written as

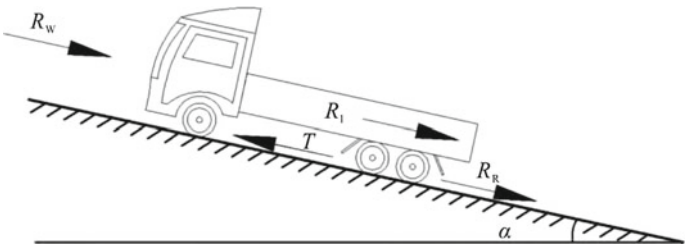


Fig. 1 Stress analysis of truck on the slope

$$D = \frac{UM_e\gamma\eta_T}{Gr} - \frac{C_D A u_a^2}{21.15G} \tag{2}$$

where U is vehicle output power ratio, which is about 90%; γ is total transmission ratio; η_T is mechanical efficiency; G is gross vehicle weight in N; r is tire radius; C_D is air resistance coefficient; A is vehicle frontal area; u_a is vehicle driving speed; M_e is engine torque.

2.2 Stable Slope Gradient and Slope Length

Vehicle driving speed on uphill will reach to a stable condition after a while. According to vehicle dynamics, vehicle dynamic factor could be written as

$$D = \frac{T - R_w}{G} \tag{3}$$

Equation (1) also could be expressed by

$$D = f + i + \frac{\delta}{g}\frac{dv}{dt} \tag{4}$$

where f is rolling resistance coefficient, i is slope gradient, δ is correction coefficient of vehicle rotating mass.

While $\frac{dv}{dt} = 0$, D could be expressed by

$$D = \psi = f + i \tag{5}$$

where ψ is road adhesion coefficient, g is gravity acceleration, whose value is 9.8 m/s^2.

By Eq. (6), the stable driving speed on uphill could be obtained

$$u_p = \frac{-Q - \sqrt{Q^2 - 4P(W - f - i)}}{2P} \tag{6}$$

where u_p is vehicle stable driving speed on uphill; P, Q, W are vehicle performance parameter.

Vehicle has the maximum dynamic factor D_{Imax} on the first gear. Under this condition, vehicle has the maximum climbing degree $\alpha_{\max}$. The maximum dynamic factor could be written as

$$D_{\text{Imax}} = f\cos\alpha_{\max} + \sin\alpha_{\max} + \frac{\delta}{g}\frac{dv}{dt} \tag{7}$$

While $\frac{dv}{dt} = 0$, Eq. (7) will be expressed by

$$D_{\text{Imax}} = f\cos\alpha_{\max} + \sin\alpha_{\max} \tag{8}$$

Therefore, vehicle's maximum climbing degree could be obtained by

$$\sin\alpha_{\max} = \arcsin\frac{D_{\text{Imax}} - f\sqrt{1 - D_{\text{Imax}}^2 + f^2}}{1 + f^2} \tag{9}$$

Vehicle driving speed will reach to a stable condition on uphill after a will. Given the original driving speed of V_1 and the stable driving speed of V_2, the driving distance, or the length of slope could be obtained by

$$S = \frac{1}{12.96}\int_{v_1}^{v_2}\frac{v}{a}\,dv = \frac{1}{12.96}\frac{\delta}{g}\int_{v_1}^{v_2}\frac{v}{\lambda D - \psi}\,dv \tag{10}$$

where a is braking deceleration.

2.3 Driving Features on Longitudinal Slope

2.3.1 Uphill Characteristics

Due to the differences on engine power and load rating of trucks, their driving status differs much. Generally, trucks have more velocity reduction when they drive the same distance on slopes with bigger slope gradient. With the slope length increases, truck driving speed declines at the first and then almost keep constant (Fig. 2). As for different trucks, vehicle has more velocity reduction on the condition of the same driving distance when it has smaller output power ratio (Fig. 3).

The great speed reduction of trucks on longitudinal slopes would result in sever traffic jam. Furthermore, heavy trucks have limited climbing performance. While the slope gradient is too large, heavy trucks could not pass the slope section, which may lead to collision incident. While passenger cars are affected slightly on slopes, frequent lane changing operation are much more universe on them, which will also affect heavy trucks. Their hence, it is of great significance to study the optimum design of slope gradient and slope length so as to improve trucks driving behavior on slopes.

2.3.2 Downhill Characteristics

Continuous braking is necessary when trucks drives downhill. Long time rubs of the brake system will produce much heat. While the great amount of heat cannot release in time, the temperature of brake disc will rise continually. With the increase of brake

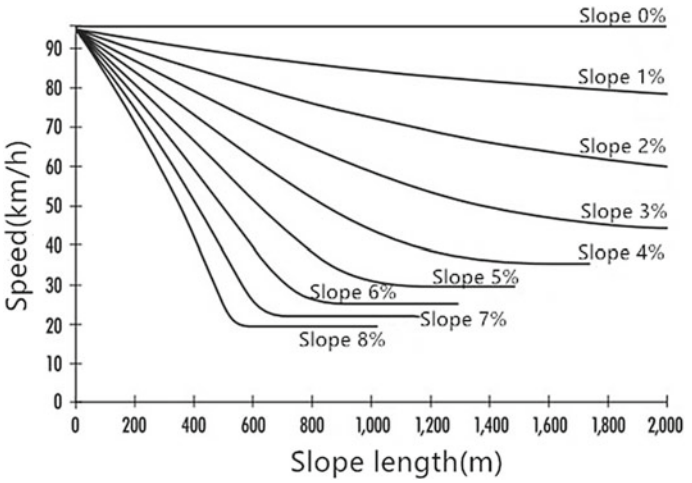


Fig. 2 The speed change of truck at different uphill

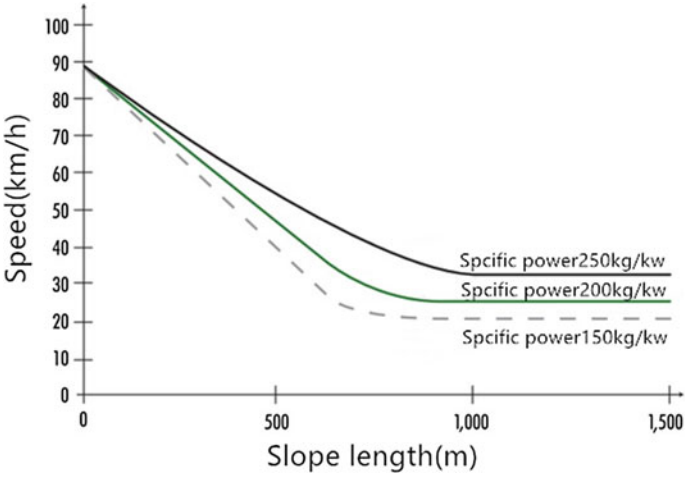


Fig. 3 The speed change of truck with different specific power climbing slope

disc temperature, the friction performance between gasket and drum decreases a lot, resulting in brake heat recession phenomenon. Severe heat recession will result in brake fade.

Factors influence disc temperature include slope gradient, gross vehicle weight, vehicle driving speed on downhill, engine braking power, the original temperature of drum, drum performance, environment temperature, and braking strategy, et al. Given the consideration that slope gradient and gross vehicle weight are the most important external factors, a great effort will be paid to investigate their influence on

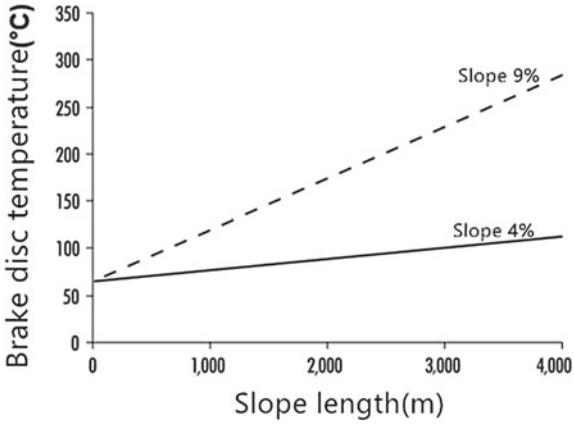


Fig. 4 The temperature change of brake disc of truck at different slope

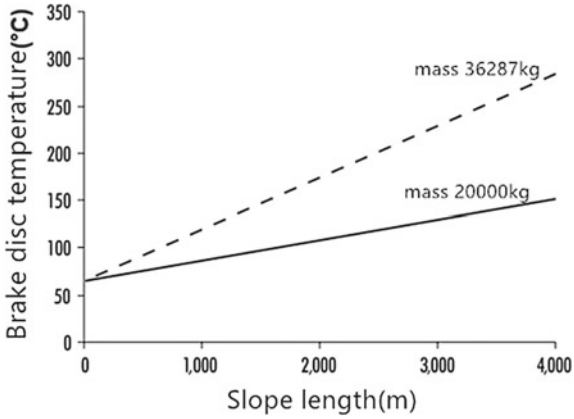


Fig. 5 The temperature change of brake disc of truck with different mass at slope

brake disc temperature. Generally, brake disc has a quicker and larger temperature rise while slope gradient is larger (Fig. 4) or gross vehicle weight is heavier (Fig. 5).

2.4 Truck Test Conditions in TruckSim Environment

According to the above analysis, truck's climbing ability depends on its output power ratio, and gross vehicle weight has a great impact on the temperature rise of brake disc. Slope gradient and slope length are two of the most important slope design parameters, and vehicle speed limits on slope with different gradients are different.

Generally, speed reduction and he temperature rise of brake disc are picked out to evaluate the influence of slope design parameters on truck climbing ability uphill and truck safe driving downhill, respectively. Therefore, TruckSim is used to investigate the speed reduction of trucks with different output power ratio and the temperature

rise of brake disc of trucks with different gross weight under the condition of different original driving speed and slope gradient.

2.4.1 The Selection of Truck Basic Parameters

Trucks may be unladen, full-loaded or partial loaded while drives uphill. It was discovered by market survey that the output power ratio of truck is about 4.5 kW/t when full-loaded and is about 9.3 kW/t when unladen. Therefore, the value of 6 and 8 kW/t were selected to describe vehicle output power ratio when vehicle is partial loaded. In the end, the value of 4.5, 6, 8, 9.3 kW/t were used to illustrate truck's output power ratio while drives uphill.

According to the limits on truck gross weight, the maximum total mass of trucks with different axles were selected as the practical vehicle gross weight. Under this condition, brake disc has the largest temperature rise while drives downhill. Therefore, 27t, 36t, 43t, 49t were used to illustrate truck's gross weight while drivers downhill.

2.4.2 The Selection of Longitudinal Slope Parameters

The design of urban underground passage should obey the City Road Design Standard of China. The regulations on the maximum slope gradient and slope length in the City Road Design Standard were listed in Tables 1 and 2. As for truck is the study subject, 80, 60, 40 km/h were used as vehicle original driving speed downhill or uphill. 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6% were used as slope gradient in simulation.

Table 1 Maximum longitudinal slope

Design speed (km/h)		100	80	60	50	40	30	20
Maximum longitudinal slope (%)	Normal value	3	4	5	5.5	6	7	8
	Limit value	4	5	6	7		8	

CJJ37-2012 "Code for Design of Urban Road Engineering"

Table 2 Maximum slope length of longitudinal slope

Design speed (km/h)	100	80	60			50			40		
Longitudinal slope (%)	4	5	6	6.5	7	6	6.5	7	6.5	7	8
Maximum slope length (m)	700	600	400	350	300	350	300	250	300	250	200

CJJ37-2012 "Code for Design of Urban Road Engineering"

Table 3 Truck uphill and downhill simulation conditions

Initial speed (km/h)	Specific power (kW/t)	Uphill slope (%)	Mass (t)	Downhill slope (%)
80	4.5	2.5	3 axles, 27 t	−2.5
		3.0		−3.0
	6.0	3.5	4 axles, 36 t	−3.5
60		4.0		−4.0
	8.0	4.5	5 axles, 43 t	−4.5
40		5.0		−5.0
	9.3	5.5	6 axles, 49 t	−5.5
		6.0		−6.0

2.4.3 Truck Typical Test Conditions

According to the above analysis, TruckSim was used to simulate vehicle drives uphill or downhill under the condition listed in Table 3.

The simulation on vehicle driving status on slopes will reveal the influence of vehicle output power ratio and the original driving speed on speed reduction uphill and that of vehicle gross weight and the original driving speed on brake disc temperature rise downhill. Analysis on the simulation result will provide data support on slope design.

3 The Establishment of Truck Simulation Model in TruckSim

3.1 Truck Models in TruckSim

3.1.1 Truck Models with the Same Specific Power and Different Axle Number

Different kinds of trucks could be established in TruckSim environment. At the first, trucks with different axle number but the same output power ratio were established to study the influence of axle numbers on vehicle uphill driving status. Basic information of trucks with output power ratio of 6.0 kW/t and different axle numbers were listed in Table 4 (Fig. 6).

Under the same original driving speed and slope gradient, speed reduction of trucks in Table 4 were plotted in Fig. 7. The stable uphill driving speed were listed in Table 5. It was concluded that vehicle axle number has little difference on its speed reduction while drives uphill.

Table 4 The basic information of trucks with different axles

	3-axle truck	4-axle truck	5-axle truck	6-axle truck
Engine maximum power (kW)	175	330	300	300
Maximum engine torque (N m)	900	5000	2000	1600
Vehicle driving mode	6 × 2, Axles 2	6 × 4, Axles 2 and 3	6 × 4, Axles 2 and 3	6 × 4, Axles 2 and 3
Total vehicle mass (t)	29	55	50	50
Tractor	–	6310 kg	6310 kg	6310 kg
Gear	Forward gear: 7, Reverse gear: 1	Forward gear: 18, Reverse gear: 1	Forward gear: 10, Reverse gear: 1	Forward gear: 10, Reverse gear: 1
Tire	510 mm, 3000 kg Single wheel	510 mm, 3000 kg Double wheel	510 mm, 3000 kg Double wheel	510 mm, 3000 kg Double wheel

(a) 3-axle truck simulation model

(b) 4-axle truck simulation model

(c) 5-axle truck simulation model

(d) 6-axle truck simulation model

Fig. 6 Vehicle simulation dynamic models

3.1.2 Truck Models with Different Specific Power

According to the analysis in 3.1.1, trucks with 6 axles were selected as the typical truck type. In TruckSim, trucks with 6 axles and output power ratio of 4.5, 6, 8, 9.3 kW/t were established. Detailed information about them were listed in Table 6.

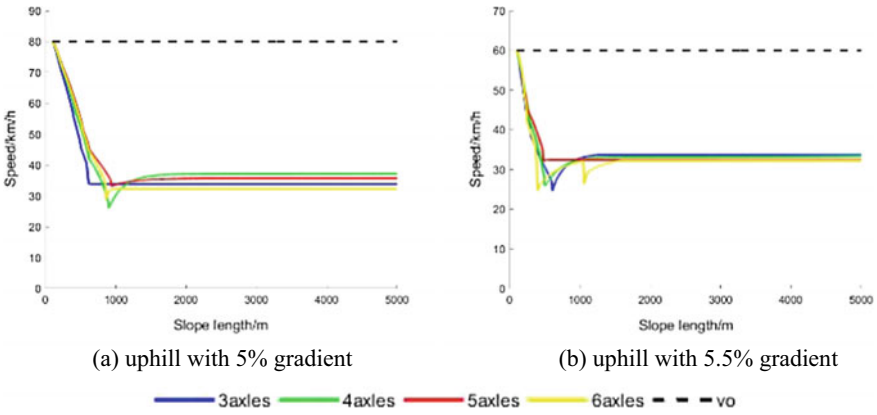


Fig. 7 The uphill speed change of truck with different number of axles

Table 5 The stable uphill speed of trucks with different number of axles

	The stable driving speed on slope with 5.0% gradient with original driving speed of 80 km/h	The stable driving speed on slope with 5.5% gradient with original driving speed of 60 km/h
Trucks with 3 axles	33.82	33.69
Trucks with 4 axles	37.21	33.36
Trucks with 5 axles	35.69	32.48
Trucks with 6 axles	32.39	32.28

Table 6 Detailed information of simulation trucks

Parameter name	Simulation parameter setting			
Number of vehicle axles	6 axles			
Number of tractor axles	3 axles			
Number of trailer axles	3 axles			
Total mass of tractor	6310 kg			
Vehicle driving mode	6 × 4, Axles 2, Axles 3			
Engine maximum power	300 kW			
Maximum engine torque	1600 N m			
Specific power (kW/t)	4.5	6.0	8.0	9.3
Total mass (t)	66	50	37	32
Gear	Forward gear: 10, Reverse gear: 1			
Tire	510 mm, 3000 kg, double wheel			

Table 7 The stable uphill driving speed of simulation and the real truck

Slope gradient (%)	The stable speed of the real truck (km/h)	The stable speed of the simulation truck (km/h)	Absolute error (km/h)	Relative error (%)
2.5	42.78	46.11	3.33	7.78
3.0	37.41	40.94	3.53	9.44
3.5	33.17	34.75	1.58	4.76
4.0	29.78	32.40	2.62	8.80
4.5	27.04	29.82	0.78	10.28
5.0	24.78	24.04	0.74	2.99
5.5	22.87	23.90	1.03	4.50
6.0	21.21	23.19	1.98	9.34

3.2 The Calibration of Truck Models with Field Tests

The calibration work of simulation trucks with the help of field tests were conducted in this research to ensure simulation accuracy. A real truck with output power ratio of 4.5 kW/t was used to carry out calibration. The stable uphill driving speed of simulation truck and the real truck under the same driving condition were listed in Table 7

In Table 6, the real truck has output power ratio of 4.46 kW/t, and the simulation truck has output ratio of 4.5 kW/t. There is a little difference on vehicle output power ratio between them. For the stable uphill driving speed, the simulation truck has slight larger stable speed than that of the real truck (the average absolute error of 2.01 km/h and the average relative error of 7.24%). The results coincide with vehicle condition. Therefore, trucks in the TruckSim environment would have a quite accurate simulation results while vehicle drives uphill or downhill.

4 Analysis of Truck Driving Behavior on Longitudinal Slopes

4.1 Truck Driving Speed Deceleration on Uphill and Optimum Design of Uphill

4.1.1 The Influence of Slope Gradient on Truck Driving Speed Deceleration

The speed reduction of trucks with different output power ratio but the same original driving speed when drives uphill were listed in Figs. 8, 9 and 10. The results showed

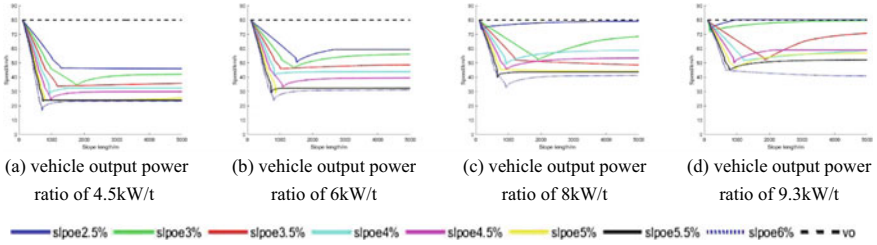


Fig. 8 The speed change of truck with initial speed 80 km/h climbing on different slope

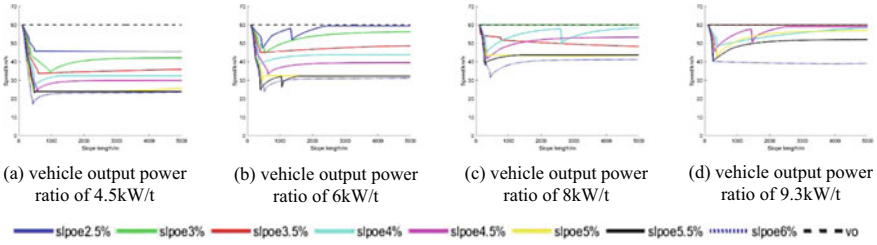


Fig. 9 The speed change of truck with initial speed 60 km/h climbing on different slope

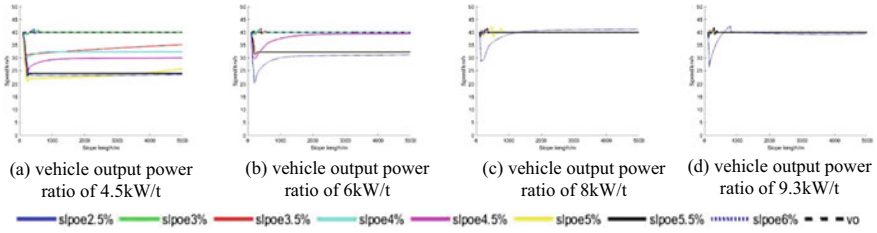


Fig. 10 The speed change of truck with initial speed 40 km/h climbing on different slope

that slope gradient has great influence on vehicle uphill speed reduction. While drives uphill, its driving speed decreases at the first, then almost keep constant. Vehicle has quicker and larger speed reduction when slope gradient is larger. Therefore, suitable slope gradient will be conductive to improve truck's uphill mobility and ensure vehicle driving safety.

4.1.2 The Influence of Vehicle Dynamic Performance on Truck Driving Speed Deceleration

The speed reduction of trucks with the same original driving speed but drives on slope with different gradient were listed in Figs. 11, 12 and 13. It was shown that vehicle output power ratio has great impact on vehicle uphill speed reduction. While

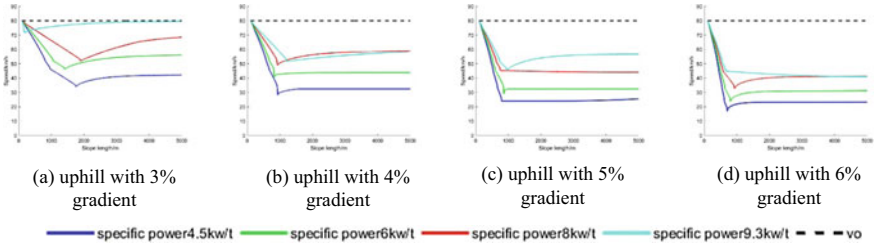


Fig. 11 The speed change of trucks with different specific power and same initial speed 80 km/h

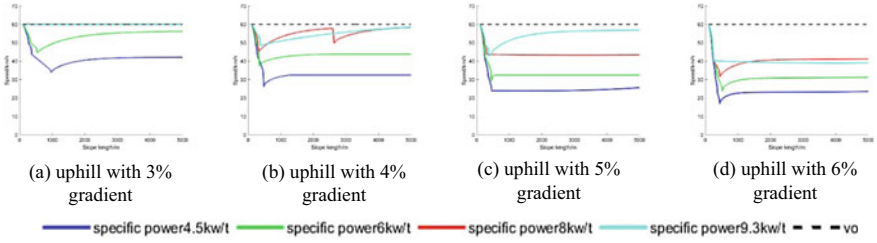


Fig. 12 The speed change of trucks with different specific power and same initial speed 60 km/h

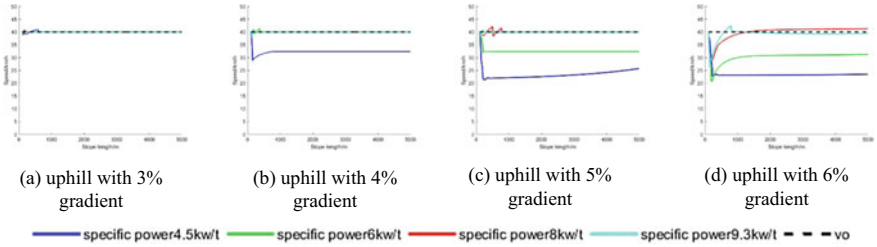


Fig. 13 The speed change of trucks with different specific power and same initial speed 40 km/h

vehicle has quite large output power ratio, its stable uphill driving speed is close to its original velocity, and the speed reduction is quite little. Given the consideration that truck type differs in different district, it is necessary to select suitable longitudinal slope design parameters according to the most universal truck's output power ratio.

4.1.3 Optimum Design for Uphill Section Based on Truck Driving Speed Deceleration

According to the analysis of vehicle uphill driving speed, it was known that the stable uphill driving speed decreases with the slope length increases. Given the consideration of vehicle dynamic performance and traffic composition in China, the

Table 8 Slope length for slope with different gradient and vehicle with different original driving speed and output power ratio

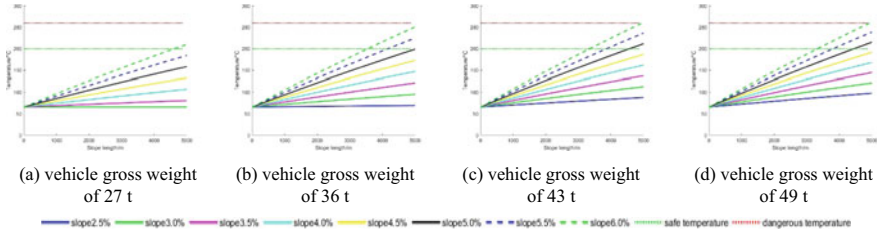
Specific power speed gradient (%)	4.5 kW/t			6.0 kW/t			8.0 kW/t			9.3 kW/t		
	80	60	40	80	60	40	80	60	40	80	60	40
6.0	326	246	—	353	275	—	399	—	—	40	—	—
5.5	358	256	—	383	289	—	442	—	—	498	—	—
5.0	381	274	—	421	327	—	501	—	—	580	—	—
4.5	418	306	—	472	—	—	585	—	—	706	—	—
4.0	468	353	—	544	—	—	713	—	—	922	—	—
3.5	538	430	—	649	—	—	936	—	—	—	—	—
3.0	642	—	—	822	—	—	—	—	—	—	—	—
2.5	810	—	—	1157	—	—	—	—	—	—	—	—

maximum slope length was defined as the slope length when the speed reduction between the original uphill driving speed and the stable velocity is 20 km/h [11]. Then, the maximum slope length for slope with different gradient and vehicle with different original driving speed and output power ratio was listed in Table 8. In this table, "—" means the maximum speed reduction could not reach to 20 km/h.

4.2 Brake Disc Temperature on Downhill and Optimum Design of Downhill

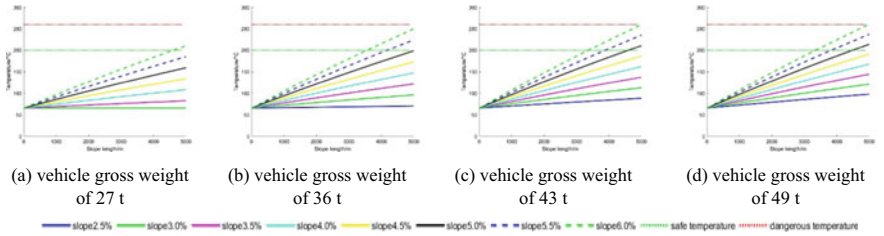
4.2.1 The Influence of Downhill Slope Gradient on Brake Disc Temperature

The temperature rise of brake disc of trucks with the same gross weight but drives on slopes with different gradient were listed in Figs. 14, 15 and 16. It was shown that the slope gradient has significant influence on temperature rise of brake disc. Brake disc has much larger temperature rise when slope gradient is bigger. Under this condition, brake system has sever heat recession, the standard braking performance could not be remained, which will degrade vehicle driving safety. There hence, it is



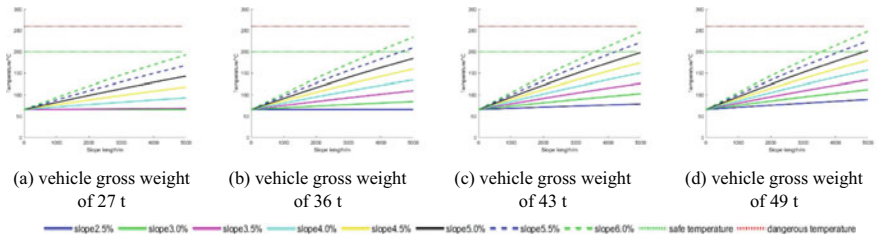
(a) vehicle gross weight of 27 t (b) vehicle gross weight of 36 t (c) vehicle gross weight of 43 t (d) vehicle gross weight of 49 t

Fig. 14 The temperature change of trucks with initial speed 80 km/h at downhill



(a) vehicle gross weight of 27 t (b) vehicle gross weight of 36 t (c) vehicle gross weight of 43 t (d) vehicle gross weight of 49 t

Fig. 15 The temperature change of trucks with initial speed 60 km/h at downhill



(a) vehicle gross weight of 27 t (b) vehicle gross weight of 36 t (c) vehicle gross weight of 43 t (d) vehicle gross weight of 49 t

Fig. 16 The temperature change of trucks with initial speed 40 km/h at downhill

quite necessary to select suitable slope gradient according to the temperature rise of brake disc, so as to get rid of sever heat recession and brake fade.

4.2.2 The Influence of Gross Truck Weight on Brake Disc Temperature

The temperature rise of brake disc of trucks with different gross weight when drives downhill were listed in Figs. 17, 18 and 19. It was shown that vehicle gross weight has significant influence on temperature rise of brake disc. Brake disc has much larger temperature rise when vehicle gross weight is larger. Given the consideration that truck type differs in different district, it is necessary to select suitable longitudinal slope design parameters according to the most universal truck's gross weight.

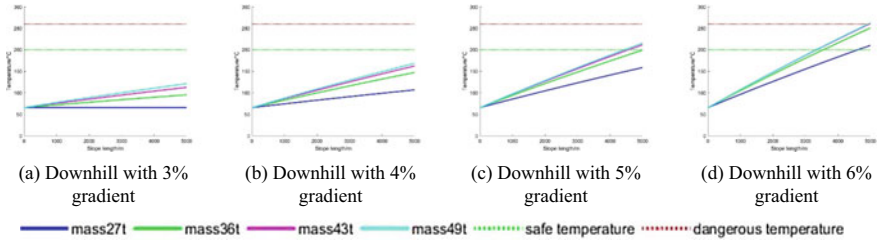


Fig. 17 The temperature change of brake disc of trucks with different mass and initial speed 80 km/h at downhill

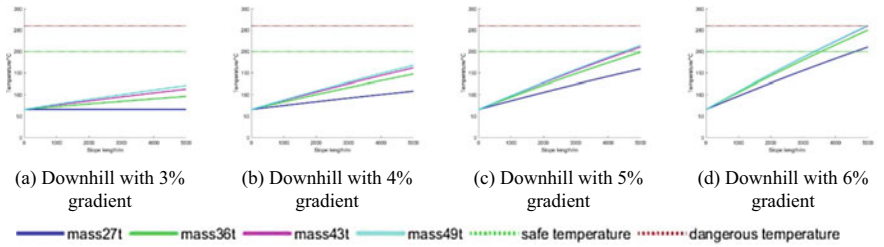


Fig. 18 The temperature change of brake disc of trucks with different mass and initial speed 60 km/h at downhill

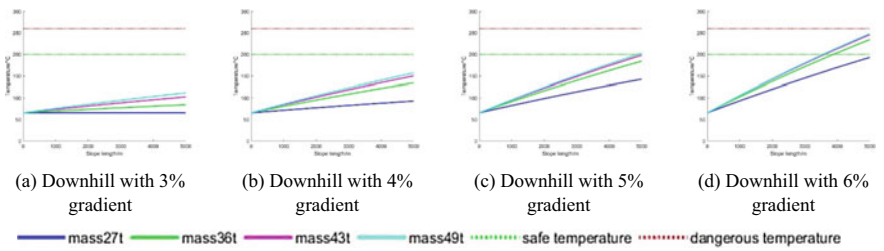


Fig. 19 The temperature change of brake disc of trucks with different mass and initial speed 40 km/h at downhill

4.2.3 Optimum Design on Downhill Section Based on Brake Disc Temperature

According to the analysis of temperature rise of brake disc, it was known that the temperature of brake disc rises continually along with drivers much frequent braking operation. While the temperature of brake disc is higher than 260 °C, brake fade comes along. To ensure the performance of brake system, the temperature of brake disc should not exceed 200 °C. Therefore, slope length when brake disc temperature reaches to 200 and 260 °C while vehicle drives downhill were obtained and shown in Tables 9 and 10. In these two tables, "—" means brake disc temperature could not

Table 9 Slope length when brake disc temperature reaches to 200 °C when drives downhill

Mass speed slope (%)	27 t			36 t			43 t			49 t		
	80	60	40	80	60	40	80	60	40	80	60	40
6.0	4650	4600	—	3550	3600	3900	3350	3400	3650	3350	3350	3600
5.5	—	—	—	4200	4200	4650	3900	3900	4250	3850	3850	4150
5.0	—	—	—	—	—	—	4600	4600	—	4450	4500	4900

Table 10 Slope length when brake disc temperature reaches to 260 °C when drives downhill

Mass speed slope (%)	27 t			36 t			43 t			49 t		
	80	60	40	80	60	40	80	60	40	80	60	40
6.0	—	—	—	—	—	—	5000	—	—	4950	5000	—
5.5	—	—	—	—	—	—	—	—	—	—	—	—
5.0	—	—	—	—	—	—	—	—	—	—	—	—

reach to 200 °C or 260 °C.

5 Conclusion

The paper puts an effort on uphill and downhill section optimum design according to trucks diving status. Typical truck models were established in TruckSim environment, the calibration work by real truck was conducted to verify the accuracy of simulation. Uphill speed reduction and downhill temperature rise of brake disc were used to obtain satisfied slope design parameters. Vehicle driving status on uphill and downhill were simulated and the results were analyzed. In the specific application, the appropriate longitudinal slope design parameters should be selected according to the condition of the most universal truck.

Acknowledgements This research was supported in part by the National Key R&D Program of China 2017YFC0804800 (2017YFC0804802), the Traffic Construction Engineering Research Project of the Shenzhen Traffic Facilities Construction Center under Contract MWKH-ky-002, and the Central Public-Interest Scientific Institution Basal Research Fund 2016-9003.

References

1. Castillo-Manzano JI, Castro-Nuño M, Fageda X (2016) Exploring the relationship between truck load capacity and traffic accidents in the European Union. Transp Res Part E 88:94–109
2. Council N (2003) Review of truck characteristics as factors in roadway design. NCHRP Report
3. Zhuang C, Zhao Y, Pan B et al (2009) Research on key parameter of highway longitudinal grade design. China J Highw Transp 22(4):39–44
4. Dong Z, Ni F, Liu S et al (2014) Heavy truck climbing speed on long and steep longitudinal slope section. J Chang'an Univ (Nat Sci) 34(3)
5. Dong Z, Lü P (2008) Influence of axle load and speed on dynamic response of semi-rigid asphalt pavement. J Chang'an Univ (Nat Sci Ed) 28(1):32–36
6. Liu H, Fu R, Zhou R et al (2004) Research report on safety and security technology for continuous long downhill sections. Chang'an University, Xi'an
7. Su B (2009) Research on continuous braking performance of large freight cars and optimization design of longitudinal slope of mountainous highway
8. Han Y, Xu J, Liu Y et al (2010) Safety slope length for driving at long-steep downgrade highway. J Chang'an Univ (Nat Sci Ed) 5:35–39
9. Wu J, Yang X, Wu L et al (2011) Risk Analysis of braking failure in continuous downhill section of Beijing National Highway G110 proposed line scheme. China Foreign Highw 1:259–262
10. Du B, Fang S, Chi S (2010) Using of truck braking in security research of long and steep downgrade on highway. J Harbin Inst Technol 42(4):656–659
11. Zhou R, Jiang L, Sun J (2004) The study of highway gradient and grade length limit. J Highw Transp Res Dev 21(7):1–4

Research on Express Highway Safety Features and Improvement Measures

Jiahui Li, Chengwu Jiao, Nale Zhao, Keman Wu and Siyuan Hao

Abstract Chinese express highway has following characteristics: incomplete control of access, divided directions, multiple lanes, heavy volume of traffic, high speed, severe lateral disturbance, etc. These lead to the highest traffic accident rate and death rate, which makes these express highways become the most unsafe highways in China. This study focuses on express highway to analysis its representative safety features, including access management, speed control and anti-dazzle facility installation, then points out improvement measures. Through optimizing road cross-section layout, controlling access density, installing traffic control facilities, the access management can be realized. To realize speed control, speed buffer zone could be installed between two adjacent speed zones with different speed limit values. It is suggested that anti-dazzle facilities should be demolished at small radius curve to eliminate obscure vision.

Keywords Express highway · Safety features · Improvement measures

1 Instruction

According to the technology guide in China, highways are divided into different grades, freeway, express highway (first-class highway), second-class highway, third class highway, forth-class highway. Express highways serve under such conditions as divided directions and multiple lanes, incomplete control of accesses. Its design speed is 100/80/60, which is similar with the design speed of freeway (120/100/80). Besides, their alignment design is similar, such as, the number of lanes should be larger than or equal to 4. Therefore, the express highway could provide high-capacity

J. Li · C. Jiao (✉) · N. Zhao · K. Wu · S. Hao
Research Institute of Highway Ministry of Transport, 8 Xitucheng Rd.,
Haidian District, Beijing, China
e-mail: cw.jiao@rioh.cn

J. Li
e-mail: happyeveryday0903@126.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_101

Table 1 Traffic accidents of different highway classification in 2014

	Freeway	Express highway	Second-class highway	Tertiary highway	Forth-class highway
Highway mileage (1000 km)	111.94	85.36	348.35	414.20	2940.99
Traffic accidents per 1000 km	76	191	98	53	5
Death toll per 1000 km	51	68	38	18	2

and high speed traffic surroundings. In fact, in plain, the express highway and freeway have no difference for facilities, except access management.

Chinese express highways are built to connect satellite cities or small cities with metropolis, to meet the heavy traffic volume and access requirements. In addition, express highway will be built when there is no adequate financial support to build freeway, or there is heavy access requirements along the highway. Therefore, in order to save money and meet the access requirements, high design speed will be chosen for express highway with no access management.

Express highway could improve efficiency and connectivity, also serve local communication preferably. However, the conditions such as incomplete control of access, mixed traffic, high density accesses, severe lateral disturbance, Express highways are facing serious problems in safety [1].

Among all classified highway, the traffic mileage of express highway is the shortest, but they are the most dangerous. As illustrated in Table 1, the traffic accident rate and death rate per 1000 km of the express highway are highest. In 2014, the traffic mileage of express highway accounts for 1.91% of the total mileage, but its number of traffic accidents accounts for 15.00% of all highway traffic accidents. Its death toll is 14.32% of all highway traffic accident deaths (The People's Republic of China Road Traffic Accident Statistics Report [2]).

This study focuses on express highway to analysis its representative safety features, then points out improvement measures to improve its safety.

2 Express Highway Safety Features

2.1 According to the Requirements to Control Accesses

It follows from its definition that on express highway, vehicles travel in divided lane and divided directions, but it is allowed to control accesses according to requirements. As a consequence, there are so many accesses along the express highway. Research has shown that in the populated areas of China, the average distance of adjacent

accesses is very short, and the large access density is 5–10 times than freeway [3]. A survey conducted on a 42 km express highway in North China Plain provide the following data. There are 248 accesses, 14 intersections, 3 alternative ramps. The average distance of accesses is about 170 m, and the minimum spacing is only 10 m. Details are illustrated in Table 2.

Especially, when the express highway passes through villages and towns, the phenomenon of mixed traffic exists. Therefore, a lot of life traffic driving into express highway from accesses, leading to complex traffic composition, serious mixed traffic, frequent traffic conflict and serious traffic safety problems. Figure 1 indicates the vehicle-pedestrian accident (bump pedestrian, crush pedestrian, etc.) deaths of different highway classifications in 2014 (The People's Republic of China Road Traffic Accident Statistics Report [2]).

Table 2 The accesses types and minimum spacing

Access types	Village internal road accesses	County road access	Factory road access	Tractor-ploughing road access	Others
Proportion (%)	51	10	10	27	2
Minimum spacing	20	20	10	10	10
Proportion of minimum spacing less than 100 m (%)	45	20	60	13	25

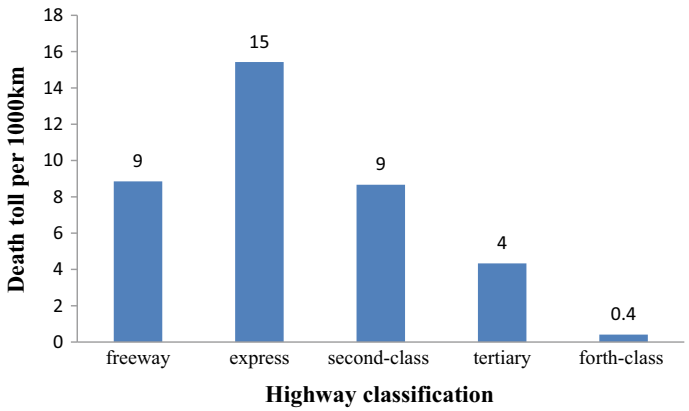


Fig. 1 Vehicle-pedestrian accident deaths of different highway classifications in 2014

As shown in Fig. 1, vehicle-pedestrian accident of express highway suffered the largest number of deaths. The design feature of express highway is similar with freeway, but its accesses could be set up according to requirements. Compared with freeway, there are so many non-motor and pedestrians get into express highway from the accesses, so it has more deaths. Moreover, compared with second-class/tertiary/forth-class highway, express highways serve under better driving condition and higher speed, so its death toll is likely to be much higher.

2.2 High Speed and Large Speed Difference

100 km/h or 80 km/h are always chosen as the design speed of express highway, which are higher than second-class/tertiary/forth-class highway. According to actual running situation, express highway has better road alignment and wider lanes, so vehicles running speed more easily reach or exceed the design speed. High speed increases the damage degree of traffic accident [3]. Researches have shown that, when speed decreased from 60 to 50 km/h, deadly accidents were reduced by 25%, injury severity score (ISS) decreased from 28 to 20 [4].

In addition, when passing through villages and towns with high speed, drivers couldn't judge the change of the surroundings and adjust their speed timely. However, the accesses disturbance leads to a lot of safety problems. In order to improve traffic safety, speed limit sign was installed before villages and towns, but the lower limit value increases the higher speed difference with the highway design speed, and the higher speed difference increases the probability of traffic accident [5]. Some studies have shown that traffic accidents are mainly caused by speed sudden change and large discreteness. According to the FHWA study results, the crash risks increases with the speed difference, are visualized in Fig. 2 [6].

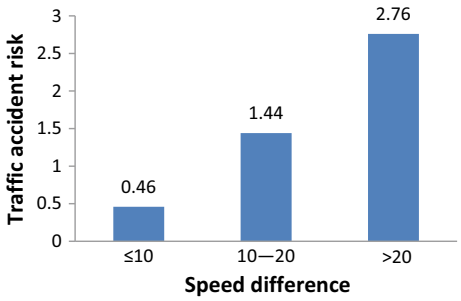


Fig. 2 The relationship between speed differences and crash risks

2.3 Contradiction Between Anti-dazzle Facilities and Sight Distance in Horizontal Curve Sections

At present, street lamp lighting is rarely used at night on the highway, and most sections of highway need automobile lighting. However, the automobile lighting will cause dazzle. Dazzle is caused by improper brightness distribution in view or extreme brightness contrast in time and space. Dazzle will lead to nausea, vision uncomfortable, short-term loss of visibility and visual fatigue [7]. In order to avoid the dazzle influence on traffic safety, Chinese design standards clearly stipulates anti-dazzle facilities installation on freeway and express highway.

Nevertheless, some of the anti-dazzle facilities' height is higher than vision of the car drivers (Guidelines for design of highway safety facilities (JTG/T D81-2006) [8]). This will affect car drivers' sight distance in small radius curve. Especially, there are a lot of openings on the medial strip when the express highway passes through villages and towns. When pedestrians get across the express highway but drivers' vision is blocked by the anti-dazzle facilities, these will be a potential security liability. Besides, the high anti-dazzle facilities lead to the situation that drivers can't catch sight of preceding barriers and vehicles in the same lane. Therefore, the anti-dazzle facilities should be optimized according to the actual situation of the express highway.

3 Safety Improvement Measures

Aim at the three problems, this paper puts forward some corresponding improvement measures.

3.1 Access Management

Express highway, as high-grade highway, is affected by its technology and higher traffic requirements, more and more accesses connected to the express highway by villages, enterprises and individuals optionally. These influence express highway traffic safety and service level severely.

Faced with above problems, this paper presents the improvement measures. That is, according to the details situation, installing dividing belt, slow traffic lane and separation facilities could be implemented to merge accesses and reduce the accesses number that connect into fast traffic lane directly (see Fig. 3). It benefits to realize scientific access management, that is, under the premise of without any increase in investment, it could provide a highly effect and quick way for vehicles to get in and out the fast traffic lane. Then it contributes to alleviate traffic jams, improve traffic safety, increase traffic capacity and reduce traffic delays.

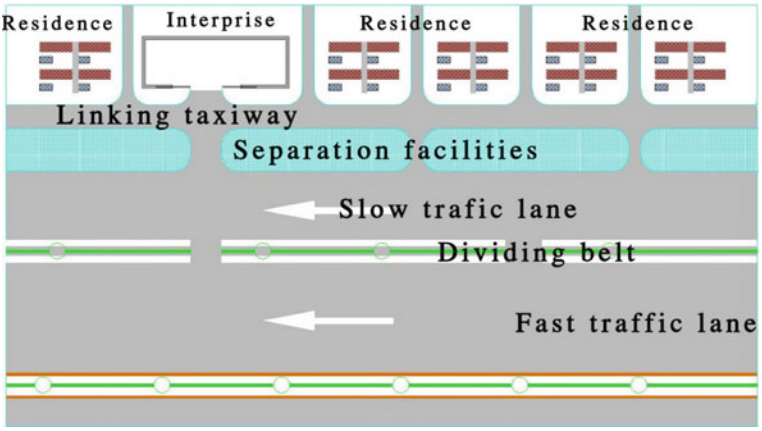


Fig. 3 Access management sketch

The above improvement measures could be realized through optimizing road cross-section layout, controlling access density, installing traffic control facilities.

3.1.1 Optimization of Road Cross-Section Layout

Most of express highways are designed to guarantee the fast traffic lane capacity, its functional attribute, access along it and lateral disturbance not taken into account. Therefore, dividing slow traffic lane, installing separation facilities and dividing belt to distinguish lane-use assignment are the main ways to optimize road cross-section layout, thus then it could reduce interfere with each other.

Separation facilities are installed to separate accesses from the fast traffic lane, thus then the accesses could connect to slow traffic lane, then dividing belts with a certain number of openings to provide passageway for vehicles from slow traffic lane to fast traffic lane. In this way, the slow traffic lane could help to realize short-range travel at the same side. Meanwhile, the openings distance is strictly controlled to provide fixed passageway for vehicles from slow traffic lane to fast traffic lane. Consequently, it could reduce the disturbance to the vehicle on the fast traffic lane.

3.1.2 Optimization of Accesses Layout

According to express highway roadside surroundings, traffic composition, crossing location, access function and central strip installation, accesses could be divided into four categories:

(1) Village internal road accesses(village internal road and express highway grade crossing);
(2) County road access(county road and express highway grade crossing);

(3) Factory road access(factory road and express highway grade crossing);
(4) Tractor-ploughing road access (tractor-ploughing road and express highway grade crossing).

Among all types of accesses, village internal road accesses along express highway are the most representative. In order to confirm the space between two adjacent accesses, the U.S. criterion- Access management Manual is quoted as the reference. According to the manual recommended method that avoid conflict overlap of right-turn traffic, when design speed is 80 km/h, the minimum space between two adjacent accesses is 120 m, that is to say, the minimum space between two adjacent openings on the separation facilities is 120 m [9].

3.1.3 Installing Traffic Control Facilities

(1) **Dividing belt**

According to the road cross-section layout optimization, if condition permission, dividing belt should be installed between slow traffic lane and fast traffic lane at the whole express highway.

(2) **Separation facilities**

For the high-density accesses, separation facilities between road and curtilage could be installed to close the express highway to reduce the adverse influence on fast traffic lanes. Separation facilities in this paper refers to the facilities that could separate vehicles and pedestrians, and then they could go on their own ways. In this way, it not only standards traffic order, but also enhance roadside pedestrians safety [10].

Separation facilities could be installed 1 m away from boundary line, then merging adjacent accesses reasonably, installing side ditch cover and taking measures to harden ground. Finally, accesses are formed. Except accesses, roadside vehicles and pedestrian can't get into the slow traffic lanes due to the presence of the separation facilities. Figure 4a, b shows sketch with and without trapezoidal ditches.

(3) **Traffic signs**

Traffic signs could provide clear, timely, enough messages for road users. Rational layout of the sign is beneficial to guide traffic and improve traffic conditions. Therefore, signs should be installed at access as shown in Fig. 5. Crossing warning sign should be installed at intersected road to indicate road right. The stop sign could help drivers stop to observe express highway traffic situation then get into express highway. This not only guarantees vehicles get into express highway safely, but also reduces the interference on the express highway at maximum extent.

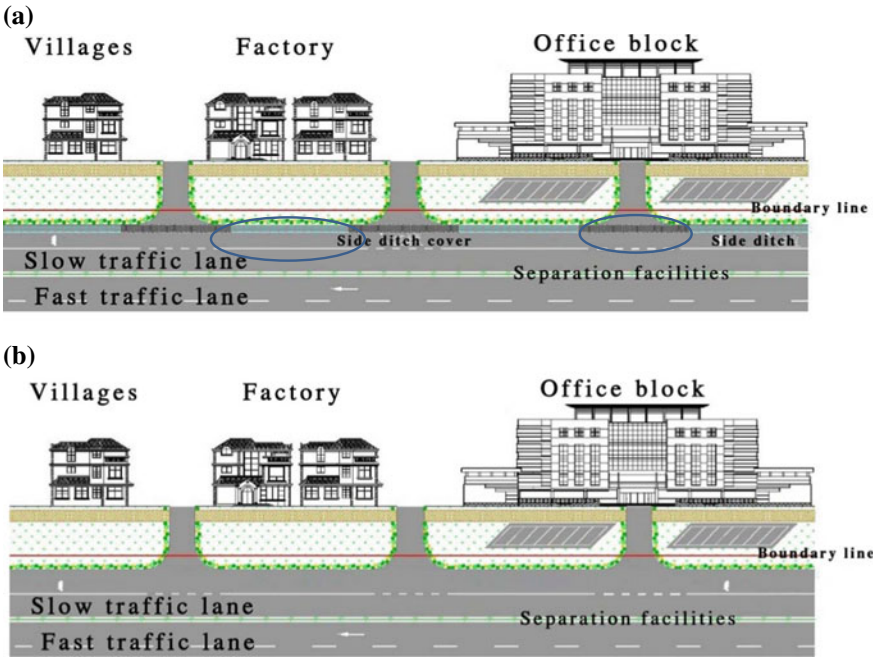
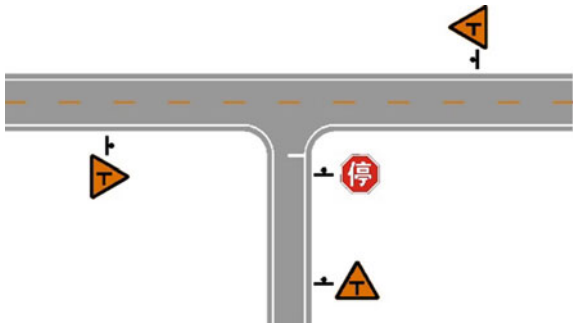


Fig. 4 **a** Express highway with trapezoidal ditches. **b** Express highway without trapezoidal ditches

Fig. 5 Traffic signs installation at access



3.2 Speed Management

Sudden change of road alignment characteristics and surroundings will cause speed change, so different speed zone should scientifically install speed limit sign to give drivers a warning about deceleration. In order to avoid large speed difference between two adjacent speed zones which causes potential traffic risks, speed buffer zone should be set up between the two speed zones.

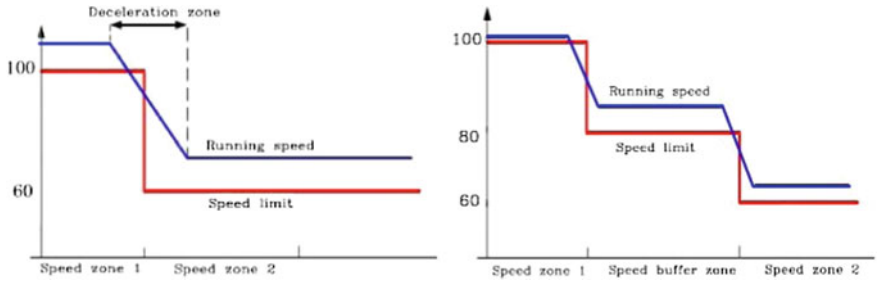


Fig. 6 Speed change comparison without and with speed buffer zone

Assuming that there is no speed buffer zone, the speed difference between two speed zone is 40 km/h, if the speed buffer zone is set up, ideally, the speed will show trapezoid decreasing trend, and descend range is about 20 km/h [5], as shown in Fig. 6.

Therefore, setting up speed buffer zone reasonably is an effective way to reduce the speed difference and improve running stationarity.

3.2.1 Speed Buffer Zone Threshold Value

According to related description about speed endurance in Chinese Guidelines for Safety Audit of Highway (JTG/T B05-2004) [11], speed difference between two adjacent speed zones is chosen as the indicators.

Assuming that vehicles running speed obeys the normal distribution, the speed difference between two adjacent speed zones also obeys the normal distribution, as shown in Formula (1).

$$\Delta V_{85} = \left| V_{85i} - V_{85(i-1)} \right| \tag{1}$$

ΔV_{85}: 85th speed difference between two speed zones;
V_{85i}: 85th speed of speed zone i;
$V_{85(i-1)}$: 85th speed of speed zone $i-1$;

If $\Delta V_{85} \leq 10$ km/h, speed endurance is good; If 10 km/h $< \Delta V_{85} <$ 20 km/h, speed endurance is medium,some measures should be taken to lower ΔV_{85} to 10 km/h; If $\Delta V_{85} \geq 20$ km/h, speed endurance is undesirable, some measures must be taken to lower ΔV_{85}.

3.2.2 Minimum Length of Speed Zone

The length of speed zone includes speed limit sign's prepositive distance and stable running distance, as shown in Formula (2).

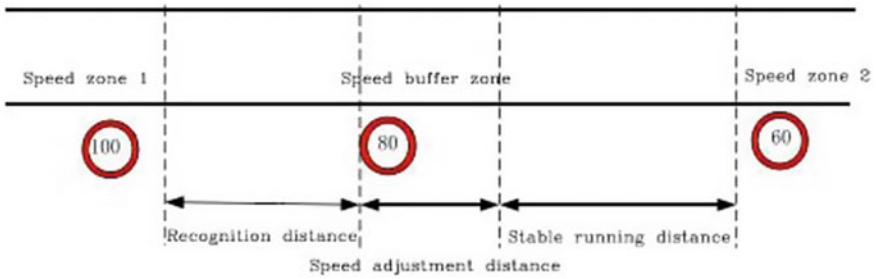


Fig. 7 Speed buffer zone length

$$L = \frac{vt}{3.6} + x \tag{2}$$

L: length of speed zone (m);
v: vehicles running speed (km/h);
x: speed limit sign's prepositive distance.

According to a previous research [5], when speed limit value is less than 80 km/h, the stable running time is 40 s. Therefore, stable running distance is about 890 m. Speed limit sign's prepositive distance could calculate on the basis of signs comprehension process.

3.2.3 The Length of Speed Buffer Zone

The length of speed buffer zone includes recognition distance S, speed adjustment distance L_1, stable running distance L_2, as shown in Fig. 7.

(1) **Recognition distance**

Recognition distance includes the distance to read information, decision distance, reaction distance and part of action distance as shown in Fig. 8 and the calculating method is shown in Formula (3).

$$s = b + c + j + d/\tan\theta = v * (t_1 + t_2 + t_3) + d/\tan\theta \tag{3}$$

s: Recognition distance; d: vertical height difference between sign and driver line of sight
t_1: Reading time, 1.5 s; t_2: Decision time, 2 s; t_3: Reaction time, 1.5 s.

(2) **Speed adjustment distance**

The speed adjustment distance from high speed zone to low speed zone computation formula is as follows (Guidelines for Safety Audit of Highway (JTG/T B05-2004) [11]):

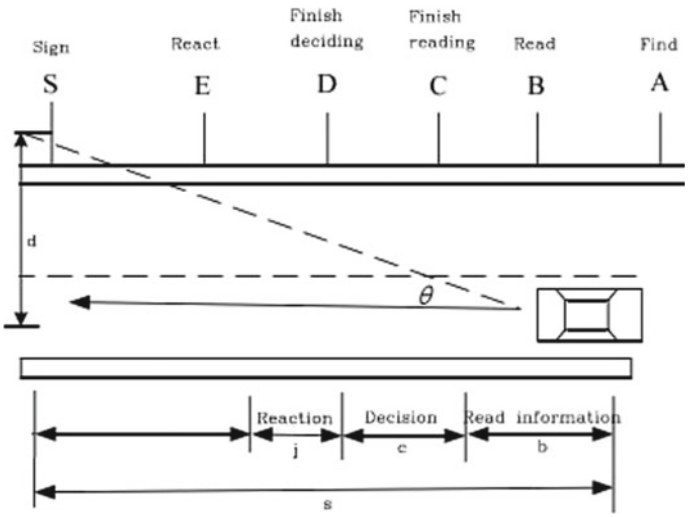


Fig. 8 Recognition distance

$$L_2 = \frac{v_1 t}{3.6} + \frac{v_1^2 - v_2^2}{2g(\varphi \pm i) * 3.6^2} \tag{4}$$

v_1: Speed limit value of speed zone 1, km/h;
v_2: Speed limit value of speed zone 2, km/h;
t: Driver's response time, 2.5 s;
φ: Longitudinal friction coefficient, 0.13;
i: Road longitudinal slope;
g: Gravitational acceleration, 9.8 m/s^2.

Therefore, the speed adjustment distance could be calculated according to the two adjacent speed zones.

(3) **Stable running distance**

According to the AASHTO related design standards (Geometric design of highways and streets), drivers need 6–10 s to judge and process surroundings information, 4–4.5 s operation time.

$$L_3 = \frac{vt}{3.6} \tag{5}$$

According to the response time and operation time, L_3 could be calculated.

3.3 Anti-dazzle Facilities Installation

Affected by visual condition, traffic accidents at night is different from daytime. According to our field study, the traffic accidents time-distribution regularity is analyzed. 19:00 pm–4:00 am is regarded as nighttime, drivers need to depend on automobile lighting during this time. The analysis results indicated that the most traffic accidents happened during daytime. Besides, there are few traffic volume and few lateral disturbance in nighttime on the express highway. Therefore, when express highway passes through villages and towns, the lateral disturbance during daytime as the main influence factor.

Vision of the car drivers will be blocked by anti-dazzle facilities on the inner lane of small radius curves, as shown in Fig. 9. This leads to drivers can't catch sight of preceding barriers and vehicles in the same lane. Therefore, synthesizes the lateral clearance, cross-sectional parameters and stopping sight distance and curve radius can be calculated using Formula (6) [12].

$$LC = R \times \left(1 - \cos \frac{90 \times S}{\pi \times R}\right) \tag{6}$$

LC: lateral clearance;
R: curve radius;
S: stopping sight distance.

According to Chinese Technical Standard of Highway Engineering (JTG D20-2006) [13], if the express speed is 80 km/h, stopping sight distance should be 110 m. Assuming that the width of inner lane is 3.75 m and the width of middle strip is 2 m, the lateral clearance should be 2.875 m. Therefore, the minimum radius is 525 m if anti-dazzle facilities will not affect vision of the car drivers. That is to say, for this road condition, curve radius is less than 525 m, anti-dazzle facilities are suggested to be demolished.

In order to improve traffic safety at night after demolishing anti-dazzle facilities, colored pavement markings are suggested to be paved before the curved section to control running speed.

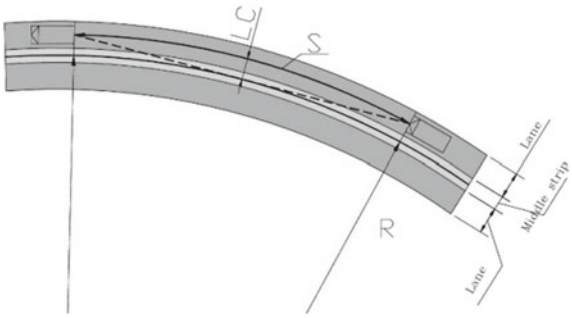


Fig. 9 Vision conditions and anti-dazzle facilities on small radius curves

4 Conclusion

In this paper, researchers discuss the Chinese-specific highway-express highway. Affected by its technical characteristics, the express highway has some representative problems, such as, access management, speed control and anti-dazzle facilities installation. For the three problems, combined with our field study, the corresponding improvement measures are presented. The main ideas are as follows:

(1) For the access management, due to a large number of accesses connect to express highway optionally, the express highway is affected by lateral disturbance seriously. Therefore, in order to reduce the lateral disturbance, researchers suggest that access management could be realized through optimizing road cross-section layout, controlling access density, installing traffic control facilities;
(2) For the speed control, with better road alignment and wider lanes, the vehicles speed on express highway is high. Besides, when express highway passes villages and towns, some speed limit signs installation lead to large speed difference. In order to reduce speed and speed difference reasonably, it is suggested that speed buffer zone should be set up between two adjacent speed zones with different speed limit values;
(3) For the anti-dazzle facility which is used to prevent glare from subtended vehicles at night, according our field study, the height of anti-facilities is higher than vision of the car drivers. It will affect car drivers' sight distance in small radius curve sections and lead to the fact that drivers can't catch sight of preceding barriers and vehicles in the same lane. Traffic accidents and traffic volume at night is much less than daytime; besides, a large number of lateral disturbance during daytime. Researchers suggest that anti-dazzle facilities should be demolished at small radius curve to eliminate obscure vision and meanwhile paving colored pavement markings before the curved section to control running speed at night.

In this study, some representative safety problems have been discussed and the corresponding improvement measures have been proposed. Among these suggested measures, some have been applied in engineering practices, but the safety benefits of them have not been studied in detail. Therefore, in the future study, it is desirable to select some express highways (which has used the above measures) to evaluate the effectiveness of the suggested measures. Besides, this study is focused on the express highways' technical character, in future study, its scientific character could be studied.

Acknowledgements This research is supported by "National key research and development plan (2017YFC0840203)".

References

1. Gong XL (2009) First-class highway traffic safety analysis and evaluation of Inner Mongolia. Master's thesis of Chang'an University
2. The Ministry of Public Security Traffic Management Bureau (2014) The People's Republic of China road traffic accident statistics report
3. Yang AG, Zheng YB (2013) Development analysis and planning guidance of Chinese express highway. Comprehensive transportation, pp 16–20
4. Wang YH, Li FCH et al (2007) Speed limit and traffic safety. J Jilin Inst Architect Civ Eng Inst 24:15–18
5. Xu T (2011) Speed zone division and speed transition zone setting research. Doctoral dissertation of Beijing University of Technology
6. Liu YT (2004) Manual of traffic safety. China Communications Press, Beijing, pp 72–74
7. Li J (2002) Traffic engineering. China Communications Press, Beijing
8. Ministry of Transportation of the People's Republic of China (2006) Guidelines for design of highway safety facilities (JTG/T D81-2006). China Communications Press, Beijing
9. Guo XC, Lu GM, Ran JY et al (2008). A study on minimum access spacing of first grade highway. Highway, pp 12–16
10. Wang SG, Gao HL (2008) Design manual of highway safety enhancement project in plain area. China Communications Press, Beijing
11. Ministry of Transportation of the People's Republic of China (2004) Guidelines for safety audit of highway (JTG/T B05-2004). China Communications Press, Beijing
12. World Road Association (PIARC) Technical Committee on Road Safety (2003) Road safety manual. Route Market, Paris
13. Ministry of Transportation of the People's Republic of China (2006) Technical Standard of Highway Engineering (JTG D20-2006). China Communications Press, Beijing

Failure Propagation Analysis of Complex System Based on Multiple Potential Field

Yong Fu, Yong Qin, Lin-Lin Kou, Dian Liu and Li-Min Jia

Abstract In consideration of the complex system structure and its functional behavior, a method of analyzing the system failure propagation process based on multiple potential field model is proposed, for the sake of seeking out all the possible failure propagation paths with their lengths if faults occur. Firstly, the structure and functional behavior of the complex system is introduced based on the complex network model. Secondly, system failure properties are analyzed and the whole process of system propagation is simulated based on the proposed failure propagation model. Finally, a case study based on railway train bogie system has been implemented to demonstrate the proposed method, which shows that the proposed model and method work well on the complex system.

Keywords Reliability analysis · Failure propagation · Complex network · Bogie system · Railway train bogie system

1 Introduction

Large facilities are the lifeline of the national economy and security. With the advancement of science and technology, large facilities regarded as the complex systems, is developing toward large-scale, automation, integration with mechanic, electric, information and computer technology. Railway trains are a large social convenience facility assembled with thousands of components and component interconnections, their reliability performances are gaining an increasing concern [1].

Y. Fu · Y. Qin (✉) · L.-L. Kou · D. Liu · L.-M. Jia
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: yqin@bjtu.edu.cn

Y. Fu · L.-L. Kou · D. Liu
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_102

The analysis of system failure propagation is the key to internal cause explanations of reliability decline for the complex large facility systems. System failure propagation refers to a component failure which may lead to the performance decline or function ineffectiveness of the facility system [2]. Currently, some researchers have made many contributions in the analysis of system failure propagation. Failure modes, effect and criticality analysis (FMECA) has been applied for the car air purifier failure propagation by Noh [3]. Jafarian and Rezvani [4] have introduced fault tree analysis (FTA) for a train derailment propagation and evaluated the whole system performance. Zheng et al. [5] have introduced Bayesian network theory combined with FTA to analyze the railway train bogie system failure propagation and ranked the most important component. However, these methods combine historical fault data with expert experience for the qualitative or semi-quantitative analysis of system failure propagation, which can cause some certain errors and affect the correctness of the results. Some researchers have also proposed the system failure propagation models with continuous characteristics. Continuous-time Markov chain models [6], cellular automaton models [7] and finite element analysis models [8] have been introduced in the applications the facility system failure propagation. However, these methods are more suitable for such systems with fewer components or simple structure. When the complexity of system structure increase gradually, the computational efficiency of these methods will decrease and may cause NP hare problems.

As for the complex large facility systems, researchers often adopt the complex network theory to model and analyze them, combined with the system topology structure. Wang et al. [9] have proposed a electromechanical system failure propagation model based on its topology and analyzed the whole propagation process by the function space iteration method. Wei et al. [10] have set up a cascading failure model for the power transmission network based on its network structure and analyzed its failure propagation laws. Liu et al. [11] have proposed the aeroengine failure propagation model based on the complex network model and simulated the whole propagation process according to the k-steps propagation principle. The abovementioned methods can be successful applied for the complex large facility system failure propagation matter to a certain extent. Nevertheless, it is of difficulty to determine the accurate propagation probability between system components in practice. In the related literature of virus transmission [12–14], the transmission probability between nodes can be dealt with based on the infected rate and reparation rate. However, it is also very difficult to determine the infected rate, which is usually obtained by experience and irrelevant with the node state itself.

To the best of our knowledge, the kinetic energy and potential energy between nodes can be transformed according to the law of conversion and conservation of energy [15, 16]. Therefore, we are trying to set up a complex system failure propagation model based on the node states and their multiple potential field to describe the failure propagation mechanism and process for the complex system.

In light of the above problems, this paper proposes a failure propagation model for complex system based on the multiple potential field. The rest of the paper can be organized as follows. Section 2 includes some basic concepts of the transformation of the kinetic energy and potential energy between nodes. Section 3 introduces the

system failure propagation model based on the multiple potential field. Section 4 performs a case study based on a specific railway train bogie system. The final section makes some discussions and conclusions.

2 Preliminaries

In order to solve the problem of the accurate propagation probability in the failure propagation model, the transformation of the kinetic energy and potential energy between nodes based on the law of energy transformation and conservation is discussed in this paper.

As for the transformation of the kinetic energy and potential energy between nodes, it can refer to the state of nodes, which can analogy to the potential field. In addition to being widely used in physics, the concept of field is also well-mentioned in the behavioral risk research [17, 18]. The field is a basic existence form of material, with the characteristics of energy, momentum and mass. It can transmit the interaction between materials. With regard to the complex system, the failure of a component can spread and evolve among the system according to a certain relationship with some space-time characteristics. Therefore, the failure of a system component can be analogous to the failure potential field.

The field is generated by the field source, while the failure potential field is generated by the system failure component. The field can be expressed as the space-time function according to a certain relationship. The nearer the field source is, the stronger the energy will be. Concerning the vector field, it is directional. But for the failure potential field, the direction can be the failure propagation direction between the system components. Therefore, the failure potential field can be applied to analyze the matter of system failure propagation, which can represents the interaction between system components on the basis of component state.

According to the failure potential field theory, the potential energy E_i of the node i can be expressed as

$$E_i = m_i a_i h_i = m_i \cdot \frac{G m_j}{r_{ij}^2} \cdot r_{ij} = \frac{G m_i m_j}{r_{ij}} \tag{1}$$

where G is a constant parameter, m_i and m_j are the mass of object i and object j, respectively, a_i is the acceleration of object i, h_i is the propagation displacement of object i, R_{ij} is the distance between the object i and object j.

We define that the object l is located at the midpoint of the line between the object i and object j. The velocity of the object i can be calculated according to the law of conversion and conservation of energy.

$$\Delta E = G m_i m_l \left(\frac{1}{x} - \frac{1}{r_{ij}} \right) = \frac{1}{2} m_l \bar{v}_l^2 \tag{2}$$

where m_l is the mass of the object i, $\bar{v}_l$ is the velocity of the midpoint of the object i and object j. x is the half of the distance between the object i and object j, which can be expressed as

$$x = \frac{r_{ij}}{2} \tag{3}$$

Therefore, the velocity of the midpoint of the object i and object j can be calculated as

$$\bar{v}_l = \sqrt{\frac{2Gm_i}{r_{ij}}} \tag{4}$$

Analogy to the potential field, we can define that the failure potential field can be composed with the component a. when the component a begins to fail, this degradation performance can be propagated to component adjacent with the component a. Therefore, the average propagation velocity of the component a can be expressed as

$$\bar{v} = \sqrt{\frac{2Gm_a}{r_{ab}}} \tag{5}$$

where m_a is the failure state of the object a, r_{ab} is the distance between the component a and its adjacent component b.

3 System Failure Propagation Model Based on the Multiple Potential Field

3.1 System Structure and Functional Behavior Network Model

Large facilities are often the extremely complex electromechanical systems combined with different components and various interactions. Therefore, complex network is applied in this paper to model the structure and functional behavior of the large facility system.

We defined the large facility system structure and functional behavior network composed of n components and m connections. The components can be represented by the nodes, and the connections between each component can be represented by the edges. Figure 1 shows an example of a certain complex network with 10 nodes and 18 directed edges.

The system structure and functional behavior network model can be defined as

$$G = (V, E, R) \tag{6}$$

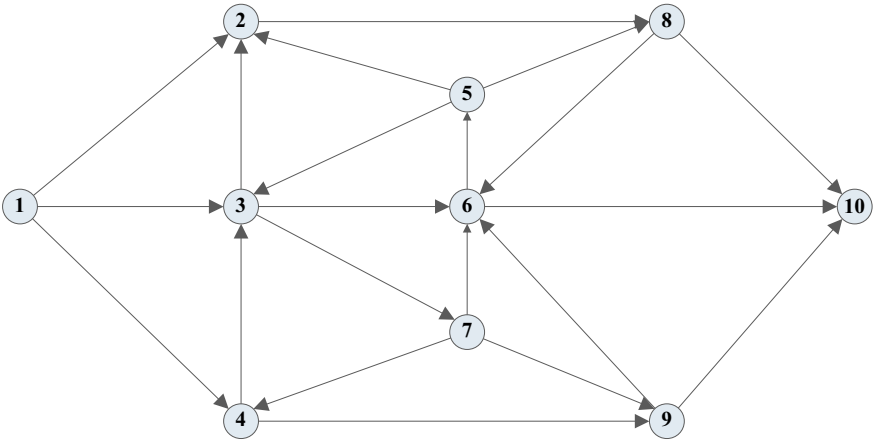


Fig. 1 Bogie system complex network model

where V represents the components of bogie system, which can be defined as $V = \{v_1, v_2, \ldots, v_n\}$. E represents the connections between each component, which can be defined as $E = \left\{e_{12}, \ldots, e_{ij}, \ldots, e_{rm}\right\}$. Besides, E can be defined as the matrix $E = [e_{ij}]$.

$$e_{ij} = \begin{cases} 1 \ (v_i, v_j) \in E \\ 0 \ otherwise \end{cases} \tag{7}$$

R represents the degree of interaction between each component, which can be seen as $R = \left\{r_{12}, \ldots, r_{ij}, \ldots, r_{rm}\right\}$. Besides, E can be defined as the matrix $R = [r_{ij}]$. The value of the r_{ij} can be obtained by the failure statistical data from the operation and maintenance department in the railway system.

3.2 Failure Propagation Model Based on the Multiple Potential Field

System failure propagation refers to a component failure which may lead to the performance decline or function ineffectiveness of its adjacent component and even affect the whole facility system. Failure propagation can trigger the failure of the successive components in the interconnected system.

Therefore, the system failure propagation model S_k can be set up on the basis of system structure and functional behavior network model $G = (V, E, R)$.

$$S_k = \langle V, E, R, M_k, W_k, I_k \rangle \tag{8}$$

where I_k represents the incidence matrix after the Kth failure propagation, which can be defined as

$$I_k = [P_{ij}]^k_{n\times n} \tag{9}$$

P_{ij} represents the probability of failure spreading from node v_i to node v_j and can be determined by the virus transmission [12–14].

$$P_{ij} = 1 - e^{-\bar{v}_{ij}\lambda_i} \tag{10}$$

λ_i is the disease duration of the node v_i, which can be defined as the mean time to resolution (MTTR) in this paper. $\bar{v}_{ij}$ is the average propagation velocity from the node v_i to node v_j, which can be determined by the failure potential field theory. However, the component v_i and component v_j can affect the propagation process. Therefore, the failure potential field model can be revised as multiple potential field model with failed node and infected node superposed simultaneously on the propagation process. We can define the transformation process of the kinetic energy and potential energy between nodes as

$$\Delta E = G\left(m_i + m_j\right)m_l\left(\frac{1}{x} - \frac{1}{r_{ij}}\right) = \frac{1}{2}m_l\bar{v}_l^2 \tag{11}$$

Therefore, the average propagation velocity from the node v_i to node v_j can be defined as

$$\bar{v}_{ij} = \bar{v}_l = \sqrt{\frac{2G\left(m_i + m_j\right)}{r_{ij}}} \tag{12}$$

M_k represents the system state after the Kth failure propagation, which can be defined as

$$\begin{aligned} M_k &= M(l)_k \\ &= M(l)_{k-1} + \left\{W(l)_{k-1} \oplus \left[M(l)_{k-1} \otimes (I_k \cdot W(l)_{k-1})\right]\right\} \end{aligned} \tag{13}$$

W_0 can be the system component affected matrix at the time of initial state of system, which can be defined as

$$W_0 = \left\{0_1 \cdots 1_i \cdots 0_n\right\} \tag{14}$$

M_0 can be the initial state of system, which can be defined as

$$M_0 = \left\{m_1^0 \cdots m_i^0 \cdots m_n^0\right\} \tag{15}$$

$M(l)_k$ is the node state set after the Kth failure propagation.

If $A \in R^{m\times n}$ and $B \in R^{1\times n}$, then the operational symbol "•" can be defined as

$$A \bullet B^T = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \bullet \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} b_1a_{11} & b_1a_{12} & \cdots & b_1a_{1n} \\ b_2a_{21} & b_2a_{22} & \cdots & b_2a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_na_{n1} & b_na_{n2} & \cdots & b_na_{nn} \end{bmatrix} \tag{16}$$

"⊗" can be defined as

$$A \otimes B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \otimes \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} b_1a_{11} & b_2a_{12} & \cdots & b_na_{1n} \\ b_1a_{21} & b_2a_{22} & \cdots & b_na_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_1a_{n1} & b_2a_{n2} & \cdots & b_na_{nn} \end{bmatrix} \tag{17}$$

"⊕" can be defined as

$$\begin{aligned} \oplus B &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \oplus \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \\ &= [\wedge(b_1, a_{11}, a_{21}, \ldots, a_{n1}) \wedge (b_2, a_{12}, a_{22}, \ldots, a_{n2}) \\ &\quad \cdots \wedge (b_n, a_{1n}, a_{2n}, \ldots, a_{nn})] \end{aligned} \tag{18}$$

When the sum of the propagation probability $\sum_k P(k)$ of the propagation path l is less than 10^{-8} [9], the whole system failure propagation can be stopped.

Therefore, the procedure of the system failure propagation based on the multiple potential field model can be explained as follows.

Step 1. Construct the system structure and functional behavior network model $G = (V, E, R)$ based on the complex theory to illustrate the system structure and its interactions.

Step 2. Determine and record the initial state of bogie system as the input of the system failure propagation model.
Step 3. Determine the failure propagation probability of each edge and find out the initial failed component as well as the component affected matrix at the time of initial state of system.
Step 4. Build up the failure propagation model centered with the initial failed component and calculate the each parameter during the iterations.
Step 5. Record the failure path along with the system failure propagation and judge whether the propagation can be stopped.

4 Case Study

In order to verify the proposed system failure propagation model based on the multiple potential field theory, this paper simulates the system failure propagation process according to an example of a specific railway train bogie system. The adopted bogie system has been built up as a system structure and functional behavior network model with 35 components, which can be seen in Fig. 2 and Table 1.

In the bogie system structure and functional behavior network, the interactions between the nodes can be the 3 connection relationships [19]: mechanical connection, electrical connection and information connection. Thus, we can transform the element "1" into "*" and "0" into "Null" in the bogie system component interaction matrix, which can be seen in the Fig. 3.

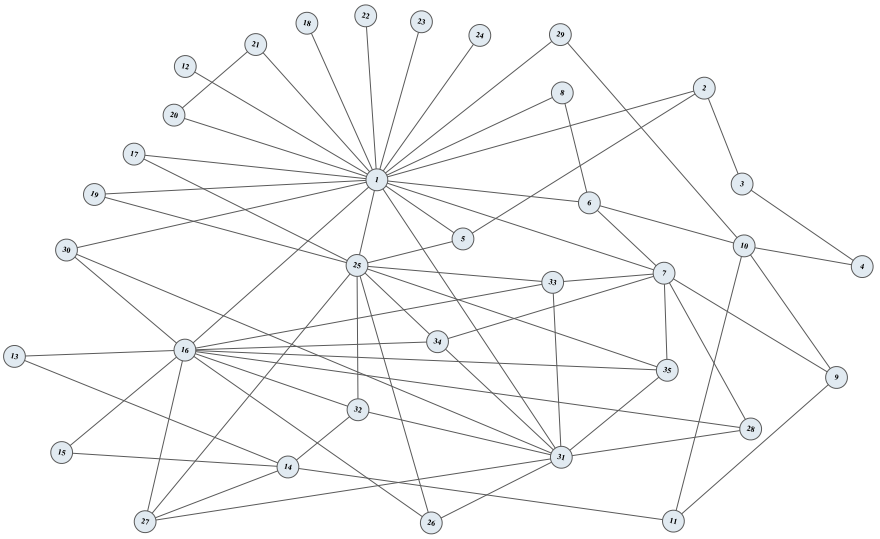


Fig. 2 Bogie system structure and functional behavior network model

Table 1 Component of a specific train bogie system

Number	Component
1	Frame assembly
2	Brake clamp
3	Brake lining
4	Brake disk
5	Pressure cylinder
6	Spring assembly
7	Axle box body
8	Primary vertical shock absorber
9	Bearing
10	Wheel
11	Axle
12	Secondary vertical shock absorber
13	Coupling
14	Gearbox assembly
15	Grounding device
16	Traction motor
17	Height adjustment device
18	Anti-hunting damper
19	Air spring
20	Center for traction pin
21	Traction rod
22	Transverse shock absorber
23	Lateral stop
24	Anti-roll bar
25	Main duct and solenoid valve
26	Velocity sensor 1
27	Velocity sensor 2
28	Velocity sensor 3
29	Surface cleaning device
30	Acceleration sensor
31	Junction box
32	Gear box bearing temperature sensor
33	Temperature sensor
34	Velocity sensor 4
35	Velocity sensor 5

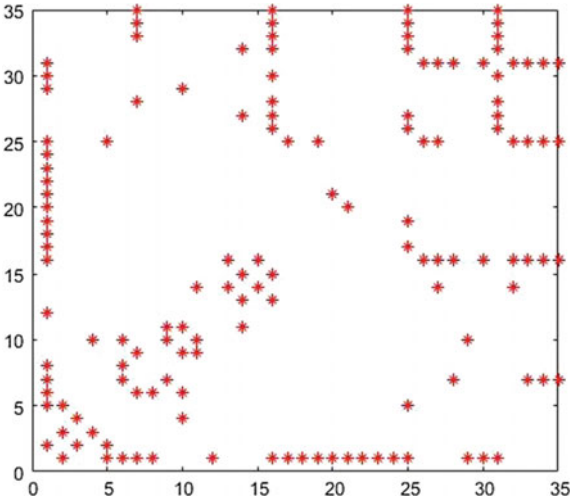


Fig. 3 Bogie system component interaction matrix

However, there may be two or three connection relationships in the system component interaction matrix, we can defined r_{ij} as [(1 with only 1 connection relationship), (2/3 with 2 connection relationships), (1/3 with 3 connection relationships)]. Therefore, the degree of interaction between bogie system components can be seen in the Fig. 4. Since there is only the mechanical connection and information connection between the component traction motor and the velocity sensor 1 at the same time, we can transform the element "1" into "*", "2/3" to "o" and "0" into "Null" in r_{ij}.

The initial state of bogie system can be determined according to the fault record data in the actual operation, which can be seen in the Fig. 5. We assume that the component 16 (traction motor) is the initial failed component and the component affected matrix at the time of initial state of system can be as follows.

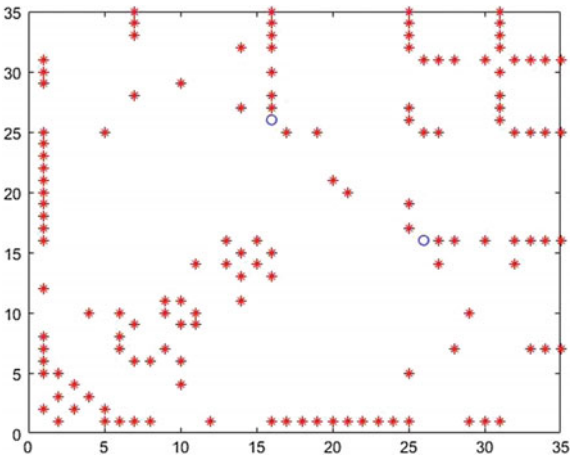


Fig. 4 Degree of interaction between bogie system components matrix

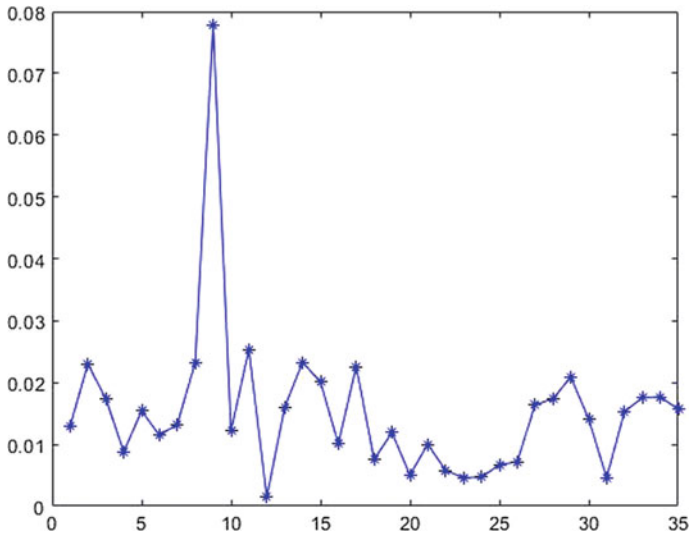


Fig. 5 Initial state of bogie system component

$$W_0 = \left(0_1 \cdots 1_{16} \cdots 0_{35}\right) \tag{19}$$

Calculate the propagation probability and construct the system incidence matrix I_1 based on the multiple potential field. The system incidence matrix I_1 can be seen in the Fig. 6, with the x-axis as the failed component, y-axis as the infected component and z-axis as the propagation probability between system components.

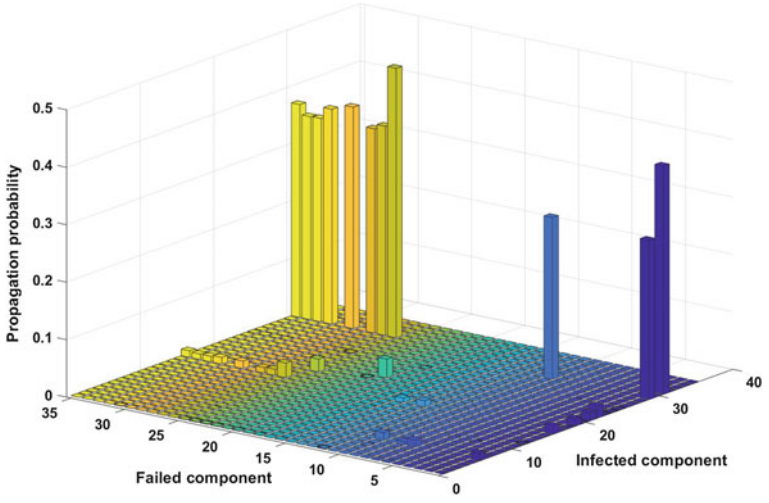


Fig. 6 Bogie system incidence matrix I_1

The failure propagation model centered with the initial failed component 16 (traction motor) can be built up based on the system initial state, the component affected matrix and the system incidence matrix I_1. Then, the failure propagation model can be iterated several times over the system state transition and the final possible paths of system failure propagation can be obtained by the terminate condition. The final possible paths of system failure propagation can be shown in Fig. 7 and Table 2.

In Fig. 7, component 16 (traction motor) fails suddenly which can be marked green, resulting in the performance deterioration of the adjacent components.

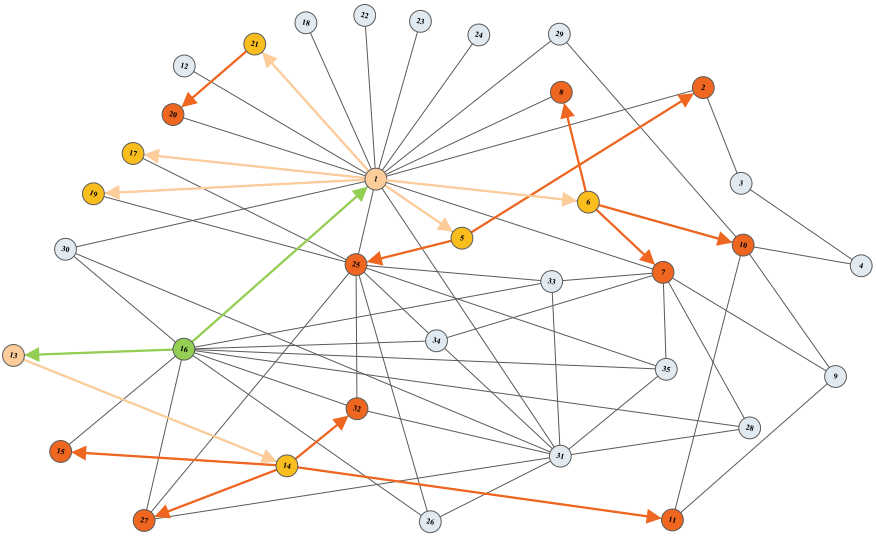


Fig. 7 Failure propagation path of bogie system with the traction motor initial failing

Table 2 Failure propagation results of train bogie system

Initial	Propagation path	Probability
v_{16}	$v_{16} \rightarrow v_1 \rightarrow v_5 \rightarrow v_2$	6.48523E−12
	$v_{16} \rightarrow v_1 \rightarrow v_5 \rightarrow v_{25}$	9.06866E−11
	$v_{16} \rightarrow v_1 \rightarrow v_6 \rightarrow v_7$	4.49397E−09
	$v_{16} \rightarrow v_1 \rightarrow v_6 \rightarrow v_8$	1.99099E−09
	$v_{16} \rightarrow v_1 \rightarrow v_6 \rightarrow v_{10}$	4.95056E−09
	$v_{16} \rightarrow v_1 \rightarrow v_{17} \rightarrow v_{25}$	2.58396E−11
	$v_{16} \rightarrow v_1 \rightarrow v_{19} \rightarrow v_{25}$	8.06498E−09
	$v_{16} \rightarrow v_1 \rightarrow v_{21} \rightarrow v_{20}$	1.18184E−08
	$v_{16} \rightarrow v_{13} \rightarrow v_{14} \rightarrow v_{11}$	1.02143E−12
	$v_{16} \rightarrow v_{13} \rightarrow v_{14} \rightarrow v_{15}$	2.00022E−12
	$v_{16} \rightarrow v_{13} \rightarrow v_{14} \rightarrow v_{27}$	3.30384E−12
	$v_{16} \rightarrow v_{13} \rightarrow v_{14} \rightarrow v_{32}$	3.85117E−12

Through simulation, it can be found that component 1 (frame assembly) and component 13 (coupling) can reach the failure state first, which can be marked yellow. At this time, the sum of the propagation probability on the failure propagation path is more than 10^{-8}, and the propagation is supposed to continue. The following simulation can be carried out. It can be found that the component 5 (pressure cylinder), component 6 (spring assembly), component 17 (height adjustment device), component 19 (air spring), component 21 (traction rod) and component 14 (gearbox assembly) first reach the failure state and can be marked orange. The propagation is also supposed to continue.

From the figure of the system reliability results, we can see that the final reliability of bogie system is no less than 0.77. At this time, the bogie system deals with some problems. This means that there must be a reliability threshold for the bogie system which can indicates whether the system is at the normal condition.

At the 3rd propagation step, component 2 (brake clamp), component 5 (pressure cylinder), component 7 (axle box body), component 8 (primary vertical shock absorber), component 10 (wheel), component 11 (axle), component 15 (grounding device), component 25 (main duct and solenoid valve), component 27 (speed sensor 2) and component 32 (gear box bearing temperature sensor) reach the failure state. At this time, the sum of the propagation probability on the failure propagation path is less than 10-8, the propagation can be stopped and the whole bogie system can be broken down.

According to the final propagation result, there are 19 components failed and bring about the system failure. Therefore, relevant measures should be taken to improve the component reliability and monitoring control should be implemented at the same time.

5 Discussions and Conclusions

Aiming at the problems of inability to describe the propagation probability between components accurately, this paper mainly discusses the application of the large facility system failure propagation analysis method and proposes a system failure propagation model based on the multiple potential field theory. The proposed model and method is demonstrated by an actual case study. This paper provides a more useful, practical and efficient way to analyze the large facility system failure propagation problem. Future work will focus on the research of the related parameter calculation precisely to improve the integrity for failure propagation models and put emphasis on the system reliability analysis based on the system failure propagation.

Acknowledgements The authors gratefully acknowledge the financial supports for this research from the State Key Program of National Natural Science of China (61833002).

References

1. Kou L, Qin Y, Jia L et al (2018) Multistate reliability evaluation of bogie on high speed railway vehicle based on the network flow theory. Int J Softw Eng Knowl Eng 28(4):431–451
2. Buldyrev SV, Parshani R, Paul G et al (2010) Catastrophic cascade of failures in interdependent networks. Nature 464(7291):1025–1028
3. Noh KW, Jun HB, Lee JH et al (2011) Module-based Failure Propagation (MFP) model for FMEA. Int J Adv Manuf Technol 55(5–8):581–600
4. Jafarian E, Rezvani MA (2012) Application of fuzzy fault tree analysis for evaluation of railway safety risks: An evaluation of root causes for passenger train derailment. Proc Inst Mech Eng Part F J Rail Rapid Transit 226(1):14–25
5. Zheng S, Qin Y, Wang L et al (2017) Bayesian network-based reliability analysis of high-speed train bogie system. In: International conference on electrical and information technologies for rail transportation. Springer, Singapore
6. Liang Z, Parlikad AK, Srinivasan R et al (2017) On fault propagation in deterioration of multi-component systems. Reliab Eng Syst Saf 162:72–80
7. Wang Z, Yao H, Yang G et al (2015) Failure risk propagation and protection schemes in coupled systems. Chaos Solitons Fractals Interdisc J Nonlinear Sci Nonequilib Complex Phenom 80:62–75
8. Chen Y, Jin Y, Liang X et al (2018) Propagation path and failure behavior analysis of cracked gears under different initial angles. Mech Syst Signal Process 110:90–109
9. Wang Y, Li M, Shi H (2018) A method of searching fault propagation paths in mechatronic systems based on MPPS model. J Central South Univ 25:2199–2218
10. Wei X, Gao S, Huang T et al (2018) Complex network based cascading faults graph for the analysis of transmission network vulnerability. IEEE Trans Ind Inform 1
11. Liu X, An S (2014) Failure propagation analysis of aircraft engine systems based on complex network. Procedia Eng 80:506–521
12. Pan Y, Yan Z (2017) The impact of multiple information on coupled awareness-epidemic dynamics in multiplex networks. Phys A Stat Mech Appl 491:45–54
13. Pan Q, Dias D (2017) Sliced inverse regression-based sparse polynomial chaos expansions for reliability analysis in high dimensions. Reliab Eng Syst Saf 167:484–493
14. Asuka N, Makoto O, Hironori Y et al (2018) The role of transforming growth factor β in cell-to-cell contact-mediated Epstein-Barr virus transmission. Front Microbiol 9:984–1005
15. Fu Y (2014) New Newton Mechanics taking law of conservation of energy as unique source law. Front Sci
16. Zborovský I (2018) A conservation law, entropy principle and quantization of fractal dimensions in hadron interactions. Int J Mod Phys A 33(10):1850057
17. Olfati-Saber R (2006) Flocking for multi-agent dynamic systems: algorithms and theory. IEEE Trans Autom Control 51(3):401–420
18. Byrne S, Naeem W, Ferguson S (2015) Improved APF strategies for dual-arm local motion planning. Trans Inst Meas Control 37(1):73–90
19. Fu Y, Qin Y, Zheng S et al (2017) Operation safety assessment of high-speed train with fuzzy group decision making method and empirical research. In: International conference on cloud computing & internet of things. IEEE

Influence of Foam Liner on Tunnels Subjected to Internal Blast Loading

Yuzhen Han, Xiuren Yang and Jingfeng Ni

Abstract The tunnels face terrorist attacks which will cause huge economic and social losses. To study the influence of geofoam liner on circular cast-iron tunnels in saturated soil subjected to internal blast loading, a new approach was introduced to simulate blast loading with LS-DYNA. The results showed that the geofoam liner could reduce the damage and provide protections to the lining due to small blast loading. Geofoam liner could be used as an effective mitigation method for tunnels.

Keywords Foam liner · Underground structures · Blast loading · Soil-tunnel interaction

1 Introduction and Literature Review

The terrorist bombing is a worldwide challenge. Metro tunnels are often the target of terrorist bombings since it will cause huge economic and social losses. Metro explosions have occurred in international metropolises such as New York, Moscow, and London. Thus, it is important and necessary to develop preventive measures to reduce the possibility of bombing attacks and avoid damage of tunnels.

Y. Han (✉) · X. Yang
Beijing Urban Construction Design & Development Group Co., Limited, No. 5, Fuchengmen Beidajie, Xicheng District, Beijing 100034, China
e-mail: hanyuzhen@bjucd.com

X. Yang
e-mail: 1006680271@qq.com

National Engineering Laboratory for Green & Safe Construction Technology in Urban Rail Transit, Beijing 100037, China

X. Yang
School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China

J. Ni
Changchun Railway Traffic Group Co., LTD, No. 999 Huaqing Road, Changchun 130021, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_103

However, there are relatively few researches about the anti-terrorist explosion of tunnel in China, and the results are rarely publicized due to the military reasons. The blast loading was not considered in the design of the ordinary civil building except for special-purpose buildings.

At present, there is no mature method for accurately describing the dynamic behavior and failure law of the segmental lining under the internal blast loading, and there are no published experimental research results.

Liu [1] found that the cast-iron subway tunnels in saturated silty soil, would suffer damage due to modest internal explosion (50–75 kg of TNT-equivalent). The tunnel vibrated obviously and transferred multiple shocks to the soil, which was also proved in Feldgun et al. [2]. Hu et al. [3] studied the damage of metro stations under internal explosion loads with numerical simulation. Feldgun et al. [2] analyzed the structural response of a circular lining under the blast load in the center, and gave a design method. Larcher et al. [4] studied the damage law of the train and the damage to passengers during the explosion.

In general, regarding the internal explosion response of the segment tunnel, there are many numerical simulation results on the response of segmental tunnel due to internal explosion [5–9]. There are few experimental results. There are fewer researches on the engineering anti-explosion measures based on the response characteristics of the segment lining [10–13].

The subject of the study is influence of foam liner on circular cast-iron tunnels subjected to internal blast loading. This research aims to improve the design and rehabilitation of transportation tunnels.

2 Numerical Analysis Model

2.1 Blast Simulation

The article simulated blast loading using a new blast loading scheme (Load Blast Enhanced) in LS-DYNA [14]. Eulerian air elements were used to model the air around the tunnel. The ambient layer works as the blast wave resource and incident blast pressure in the ambient layer are derived from CONWEP [15]. This numerical approach was verified by Han [16].

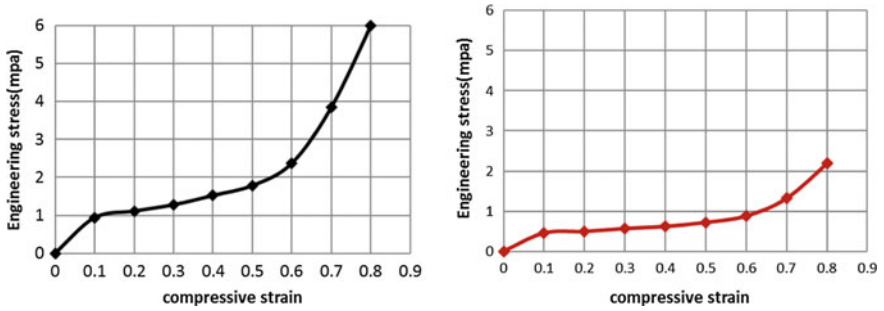
The soil was modeled with FHWA Soil Model 147. An isotropic elastic-plastic material in LSDYNA was used to simulate the cast-iron tunnel in which yielding stress versus plastic strain curves can be defined for compression and tension separately [17]. The failure tensile stress of the tunnel is 150 Mpa when the tensile strain reaches 0.004. The compressive yielding stress is 600 Mpa. The Young's modulus of tunnel is 100 Gpa. The 8-node hexahedron elements integrated using 8-point Gaussian method was used to simulated the air and ambient layer in LSDYNA. The equation of state of MAT_NULL material model was adopted. The detailed parameters were validated in Han [16].

2.2 Constitutive Model of Foam Liner

Simulation was conducted to study the effect of geofoam liner on mitigating tunnel damage. The thickness of foam layer inside the lining was 4 cm. Two kinds of geofoam were used with different densities of geofoam, 61 and 112 kg/m^3. The strain–stress characteristics of foam material [18] was shown in Fig. 1.

The simulation model with embedded foam layer in LSDYNA was shown in Fig. 2. The foam layer was 10 m in the longitudinal direction. It was quite acceptable and reasonable since there was no response of lining at the location farther than 10 m in the longitudinal direction. The mesh size of foam layer was the same as the air to transfer the blast wave accurately.

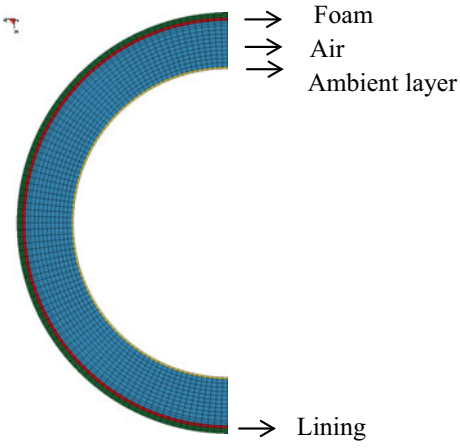
Geofoam was modeled with the crushable foam constitutive model of LSDYNA. This isotropic foam model crushes one-dimensionally with a Poisson's ratio that is essentially zero. In the implementation Young's modulus is assumed to be constant



(a) Foam material 1 (b) Foam material 2

Fig. 1 The stress-strain curve for foam material 1 and 2

Fig. 2 Simulation model with geofoam in LSDYNA (Remove the interface layer)



and the stress is updated assuming elastic behavior.

$$\sigma_{ij}^{trial} = \sigma_{ij}^{n} + E\dot{\varepsilon}_{ij}^{n+1/2}\Delta t^{n+1/2} \tag{1}$$

The magnitudes of the principal values, σ_i^{trial}, i = 1, 3 are then checked to see if the yield stress, σ_y, is exceeded and if so they are scaled back to the yield surface:

If $\sigma_y < \left|\sigma_i^{trial}\right|$, then

$$\sigma_i^{n+1} = \sigma_y \frac{\sigma_i^{trial}}{\left|\sigma_i^{trial}\right|} \tag{2}$$

After the principal values are scaled, the stress tensor is transformed back into the global system. The yield stress is a function of the natural logarithm of the relative volume, V, i.e., the volumetric strain.

3 Results of Simulations

The results showed that the tunnel had no fracture due to 25 kg TNT. The damage in the lining with and without the geofoam layer is shown in Fig. 3. The geofoam layer could reduce the damage obviously in the lining. The maximum plastic strain in the tunnel was 0.00288 with geofoam compared to 0.00296 in the case without one as shown in Table 1.

The response in the soil was also reduced obviously with the foam protections. The maximum effective shear strain in the soil with foam was 23% comparing to 31% in the case without foam protections.

Figure 4 shows the time history of acceleration at point 1 in lining with and

(a) Foam material 1 (b) Foam material 2 (c) Without foam

Fig. 3 Failure in the lining with and without embedded foam due to 25 kg TNT

Table 1 Results of simulation with and without embedded foam due to 25 kg TNT

Simulation results	Foam 1	Foam 2	Without foam
Length of damaged lining (m)	4	3.7	4.3
Max shear strain in the tunnel	0.00288	0.00288	0.00296
Max effective shear strain in the soil (%)	23	23	31

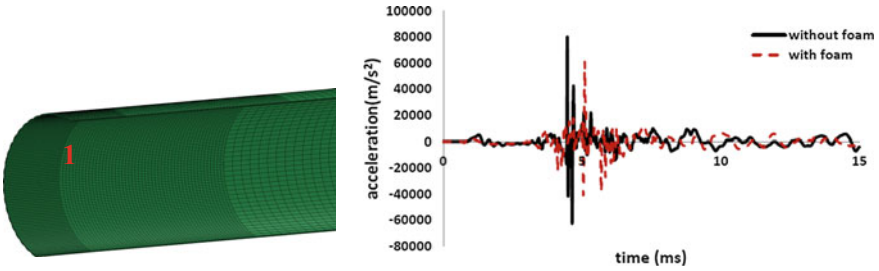


Fig. 4 Time history of acceleration at point 1 with and without foam due to 25 kg TNT

without foam protection. The magnitudes of acceleration in both expanding direction (reduced by 27.03%) and shrinking direction (reduced by 36.51%) were both reduced with the foam protection.

4 Conclusion and Future Research

- The Geofoam liner could provide protections to the lining due to small blast loading. The damage can be reduced 13.95% and the max effective shear strain in the soil can be reduced 25.81% with foam 2. The magnitudes of acceleration in both expanding direction and shrinking direction were both reduced significantly with the foam protection.
- However, it was found that with larger amount of explosive (50 kg TNT), a 4-cm liner made of general geofoam basically did not influenced the blast response of the tunnel. Future research will be conducted to study thicker geofoam liner or new material to mitigate the tunnel damage due to large internal blast loading.

References

1. Liu H (2012) Soil-structure interaction and failure of cast-iron subway tunnels subjected to medium internal blast loading. J Perform Constructed Facil ASCE 26(5):691–701
2. Feldgun VR, Kochetkov AV, Karinskia YS, Yankelevsky DZ (2008) Internal blast loading in a buried lined tunnel. Int J Impact Eng 35:172–183
3. Hu Q, Yuan, Y (2009) Numerical simulation of internal blast effects on a subway station. Springer, Berlin
4. Larcher M, Casadei F, Solomos G (2010) Risk analysis of explosions in trains by fluid-structure calculations. J Transp Secur 3(1):57–71
5. Liu H (2009) Dynamic analysis of subway structures under blast loading. Geotech Geol Eng 27(6):699–711
6. Feldguna VR, Kochetkovb AV, Karinskia YS, Yankelevsky DS (2008) Blast response of a lined cavity in a porous saturated soil. Int J Impact Eng 35:953–966

7. Chille F, Sala A, Casadei F (1998) Containment of blast phenomena in underground electrical power plants. Adv Eng Softw 29:7–12
8. Choi S, Wang J, Munfakh G, Dwyre E (2006) 3D Nonlinear blast model analysis for underground structures. In: Proceedings of geocongress 2006, Paper No. 206
9. Preece DS, Weatherby JR, Blanchat TK, Davie NT, Calderone JJ, Togami TC, Benham RA (1998) Computer and centrifuge modeling of decoupled explosions in civilian tunnels. Technical Report, Sandia National Laboratories, Albuquerque, NM
10. Choi S, Chaney I, Moon T (2010) Blast and post blast behavior of tunnels. In: Proceedings North American Tunneling Conference, Society for Mining, Metallurgy, and Exploration, Englewood, CO, USA, pp 655–663
11. He W, Chen JY, Guo J (2011) Dynamic analysis of subway station subjected to internal blast loading. J Cent South Univ Technol 18(3):917–924
12. Chakraborty T, Larcher M, Gebbeke N (2014) Performance of tunnel lining materials under internal blast loading. Int J Protective Struct 5(1):83–96
13. Lu Y, Wang Z, Chong K (2005) A comparative study of buried structure in soil subjected to blast load using 2D and 3D numerical simulations. Soil Dyn Earthq Eng 25:275–288
14. LS-DYNA (2012) Version 971 R6.0.0 Theory Manual. California, USA: Livermore Software Technology Corporation (LSTC)
15. Schwer L (2010) A brief introduction to coupling load blast enhanced with multi-material ALE: the best of both worlds for air blast simulation, German LS-DYNA Forum
16. Han Y, Liu H (2015) Finite Element simulation of medium-range blast loading using LS-DYNA. Shock and Vibration (in press)
17. Lewis BA (2004) Manual for LS-DYNA Soil Material Model 147. Technical Publication No. FHWA-HRT-04-095. Federal Highway Administration
18. Karagiozova D et al (2009) Response of flexible sandwich-type panels to blast loading. Compos Sci Technol 69(6):754–763
19. UFC 3-340-02 (2008) Structures to resist the effects of accidental explosions, Department of the Army and Defense Special Weapons Agency, Washington DC

The Optimal Road Tolls and Parking Fees for Managing Daily Household Commute in a Linear City

Yi Yao, Ling-Ling Xiao and Wei-Jiu Zhang

Abstract This paper exams a daily traffic pattern of household on a morning and evening commute link. It is assumed that the parking is located along commuting routes radiating from the CBD and the households have two preferred arrival times, i.e., school start time and work start time. Based on bottleneck model, the households' departure time choice is assumed to follow the user equilibrium principle according to tip cost. The analytical solution of the proposed model is derived. We then design pricing schemes, i.e., a time-varying road toll and a location-dependent parking fee. Within the framework of the extended bottleneck model, we proved that the proposed joint scheme of road toll and parking fee can effectively eliminate the queues behind the bottleneck, even reduce the schedule delay by reversing the spatial order of parking. We also find that the location-dependent parking fee with no road toll could improve the morning commute pattern. Furthermore, the numerical results show that the proposed pricing schemes can indeed improve the efficiency of the household commute through decreasing the total travel cost.

Keywords Household travels · Daily commute pattern · Road toll · Parking fee

1 Introduction

In most big cities, the increased demand of parking space and the dependence of commuters on private cars brings great pressure to the traffic. At the meanwhile, Household travels are common especially in Asian cities. We define household commuter as a child who needs to go to school and an adult who needs to go to CBD.

Y. Yao · L.-L. Xiao (✉) · W.-J. Zhang
School of Economics and Management, Beijing Jiaotong University, Beijing 100044, People's Republic of China
e-mail: lxiao@bjtu.edu.cn

Y. Yao
e-mail: 2803317028@qq.com

W.-J. Zhang
e-mail: 16241244@bjtu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_104

So, household travels consist of two consecutive parts, i.e. travelers drop off their children at school in the morning and then drive to CBD. But they don't need to pick up their children in the evening in our hypothesis. Although various aspects of bottleneck model and parking are researched in some studies, a comprehensive study of household travel with bottleneck model, on-street parking and bi-direction commute is still lacking. Therefore, integrated cost models are established to reflect the behavior choice of household commuters and bi-direction commuting patterns, and research the effectiveness of time-vary road toll and location-dependent parking fee.

The classical bottleneck model were put forward by Vickrey [1]. The original model was built to research the travel plans of individual commuters during the morning rush hour. It was usually assumed that the travel cost consists of two parts: the cost of travel time cost and the schedule delay cost caused by early or late arrival at CBD. Because of the bottleneck, there will be the cost of queuing time. In this case, commuters will choose departure time to avoid queuing. This process follows the user equilibrium principle, that is, all commuters will have the same commuting cost no matter when they leave home. Most subsequent studies were based on the single bottleneck model and extended hypothesis (see comprehensive reviews in Arnott et al. [2, 3], de Palma and Fosgerau [4]).

In existing literatures, commuters are most individual travelers [5, 6]. Arnott et al. [7]. Some scholars considered that travelers need to park their car near CBD, then walk through the parking area, and so on-street model was born. They built a model to analyze the travel pattern with on-street parking and single-bottleneck. Zhang et al. [8] studied daily commute that take morning and evening peaks as a whole and discussed different charging systems. In reality, a lot of commutes are household travels. Jia et al. [9] analyzed the departure pattern of household commuters and congestion fee. Liu et al. [10] study investigates the morning commute problem with both household and individual travels, where the household travel is a shared ride of household members.

In our study, we try to make a comprehensive analysis of household daily commuting. First of all, the evening rush hour commute is modeled in the case of given certain parking location. Next, we integrate the cost of evening commuting into the morning rush hour model. Finally, we introduce a time-vary road toll and location-dependent parking fee mechanism and compare their efficiency by modeling.

The paper is organized as follows. In the next Section, we first model evening travel pattern use Nash-equilibrium and then establish morning commute model without road toll and parking fee. In Sect. 3, we carry on a study at time-varying road tolls based on regime f. In Sect. 4, we propose a strategy with time-varying road tolls and location-dependent parking fees. Section 5 discusses the regime of location-dependent parking fee without road toll. Section 6 shows the numerical examples. The last Section is the conclusion of our research.

2 Morning and Evening Commuting Patterns Considering Daily Travel Cost Without Road Toll and Parking Fee (f)

In daily commute network, we consider three important places: home, school and workplace. Residents would departure from home to send their children to school, then travel to CBD for work. In the evening, commuters leave workplace and go back home without picking up their children. The traffic network is shown in Fig. 1.

In Fig. 1, we assumed the bottleneck capacity in the home-to-school-to-work direction is $\vec{s}$, and that in the other direction is $\overleftarrow{s}$. Assuming there are N household commuters need to go to school and work every day. But in the evening commute, we assume the commuters go straight home after work without picking up their children. A queue will develop if the arrival rate exceeds the road capacity. For simplicity, we assumed the free flow travel time is zero. On-street parking area is near to CBD, where the parking points are indexed by n in order of increasing distance from CBD.

After parking cars, commuters have to walk to CBD from parking location n. Total walking time is $W + wn$, where W is the walking time from the parking area to CBD and w represents walking time for passing one point. Since the cost related to W is fixed, we assume it to be zero. The travel time with car in the parking area is ignored.

The school starting time is the preferred arrival time at school $\vec{t}_1^*$, while the work starting time is the preferred arrival time at work $\vec{t}_2^*$. And the preferred leaving time from work is $\overleftarrow{t}^*$. We use $\vec{\beta}$ to represent the unit cost of early arrival time, and late arrival for school and work is not allowed. The cost of unit late leaving time is $\overleftarrow{\gamma}$, and people can't leave early. The cost of unit vehicle travel time is α, and the cost of unit walking time is λ. To satisfy the equilibrium conditions, we set that 00 and $\alpha > \overleftarrow{\gamma}$. People prefer driving rather than walking in reality, so $\lambda > \alpha$.

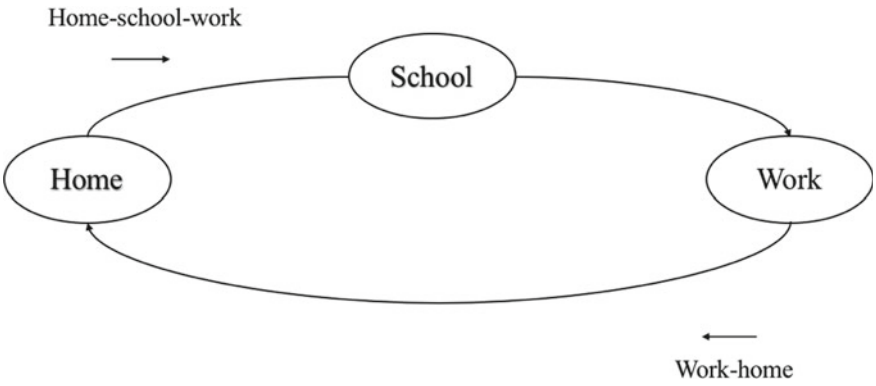


Fig. 1 A bi-direction bottleneck network

2.1 Nash-Equilibrium Travel Pattern in the Evening Commute with Fixed Parking Locations

The parking points of the evening commute is determined by the morning commute. We assume that the problem of picking up children is not taken into consideration in the evening commute, so the travel routes are regarded as individual travel. According to the research assumptions and realistic conditions, commuters departure from CBD after $\overleftarrow{t}^*$, the travel cost of commuter who leaves the bottleneck at time $\overleftarrow{t}$ with parking location n is

$$\overleftarrow{C}^f\left(\overleftarrow{t},n\right)=\alpha\frac{\overleftarrow{D}\left(\overleftarrow{t}\right)}{\overleftarrow{s}}+\lambda wn+\overleftarrow{\gamma}\left[\overleftarrow{t}-\frac{\overleftarrow{D}\left(\overleftarrow{t}\right)}{\overleftarrow{s}}-wn-\overleftarrow{t}^*\right], \tag{1}$$

where $\overleftarrow{D}\left(\overleftarrow{t}\right)$ is the length of the queue of commuter who leaving the bottleneck at time $\overleftarrow{t}$. The first part of the equation is the cost of queuing time, the second part is the cost of walking time, and the last section is the cost of late leaving from workplace. Through the derivation about $\overleftarrow{t}$, we get

$$\frac{\partial\overleftarrow{C}^f\left(\overleftarrow{t},n\right)}{\partial\overleftarrow{t}}=\overleftarrow{\gamma}+\frac{\alpha-\overleftarrow{\gamma}}{\overleftarrow{s}}\frac{d\overleftarrow{D}\left(\overleftarrow{t}\right)}{d\overleftarrow{t}}. \tag{2}$$

Since there is no queue when the evening rush hour just begin, and the length of queue can't be negative, we have $\frac{d\overleftarrow{D}(\overleftarrow{t})}{d\overleftarrow{t}}\geq 0$ during the evening rush hour. When the state of $\frac{d\overleftarrow{D}(\overleftarrow{t})}{d\overleftarrow{t}}>0$ holds for a period of time, some commuters leave early for decreasing schedule delay and queuing cost. Therefore, the travel cost increases over departure time $\overleftarrow{t}$. Because of the FIFO assumption, individual commuters hope to get into the queue as early as possible to cut travel cost. As a result, some travelers leave the office immediately after $\overleftarrow{t}^*$ because they parked cars closer to CBD. They have to walk less than others and entry the queue earlier.

In our research, we assume that $w<1/\overleftarrow{s}$ to make sure the queue forming and growing at the bottleneck. The arrival rate of bottleneck is $\overleftarrow{r}_1^f=1/w$.

But, if commuters who park their cars far from CBD prefer waiting in their offices to avoid queues, queues may decrease. Then a stable situation with constant travel costs would be reached. According to the condition $\frac{\partial\overleftarrow{C}^f(\overleftarrow{t},n)}{\partial\overleftarrow{t}}=0$, we can derive the arrival rate to the bottleneck: $\overleftarrow{r}_2^f=\frac{\alpha-\overleftarrow{\gamma}}{\alpha}\overleftarrow{s}$.

As the travel pattern depicted in Fig. 2, there is no queue for the first commuter and

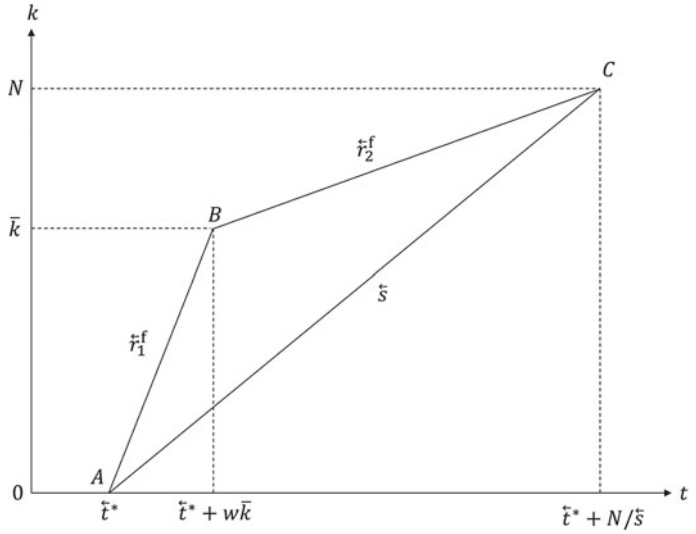


Fig. 2 Traffic pattern in the evening commute, regime f

the last commuter. In the picture, line ABC means the arrival curve to the bottleneck and AC shows the departure curve from the bottleneck. $\bar{k} = \frac{\overleftarrow{\gamma} N}{\alpha + \overleftarrow{\gamma} w \overleftarrow{s} - \alpha w \overleftarrow{s}}$ is the turning point to indicate the last commuter to leave the office immediately. Figure 2 shows arrival and departure patterns the bottleneck. The condition that commuters with closer parking spots depart earlier are suitable for $k \leq \bar{k}$. Commuters who are in the scope of $k > \bar{k}$ do not necessarily depart according to the parking order.

We define $n(k)$ as the parking location of the kth departure, we have

$$n(k)\begin{cases} = k & \forall k \in [0,\ \bar{k}]; \\ \leq \frac{N}{w\overleftarrow{s}} - \frac{\alpha(N-k)}{\left(\alpha - \overleftarrow{\gamma}\right) w \overleftarrow{s}} & \forall k \in [\bar{k},\ N]. \end{cases} \tag{3}$$

In the following analysis, we set $\bar{n} = \bar{k}$.

For the case of $n \leq \bar{n}$, $\overleftarrow{D}\left(\overleftarrow{t}\right) = \overleftarrow{s}\left(1 - w\overleftarrow{s}\right)\left(\overleftarrow{t} - \overleftarrow{t}^*\right)$, and the cost of commuter who departs from the bottleneck at time $\overleftarrow{t}$ with parking location n is

$$\overleftarrow{C}_n^f = \frac{\alpha n}{\overleftarrow{s}} - \alpha w n + \lambda w n. \tag{4}$$

For the case of $n > \bar{n}$

$$\overleftarrow{C}_n^f = \alpha \frac{\overleftarrow{D}\left(\overleftarrow{t}\right)}{\overleftarrow{s}} + \overleftarrow{\gamma}\left[\overleftarrow{t} - \frac{\overleftarrow{D}\left(\overleftarrow{t}\right)}{\overleftarrow{s}} - \overleftarrow{t}^*\right] + \left(\lambda - \overleftarrow{\gamma}\right)wn. \tag{5}$$

From the above analysis, based on user equilibrium principle, we get the equilibrium travel cost $\overleftarrow{C}^f\left(\overleftarrow{t}, n\right) = \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}$ y.

So, the evening travel cost in this basic regime for commuter with parking location n is

$$\overleftarrow{C}_n^f = \begin{cases} \frac{\alpha n}{\overleftarrow{s}} - \alpha wn + \lambda wn \quad if\ n \le \bar{n}; \\ \left(\lambda - \overleftarrow{\gamma}\right)wn + \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}} \quad if\ n > \bar{n}. \end{cases} \tag{6}$$

For commuters of $n \le \bar{n}$, they leave office as early as possible due to schedule delay will increase their travel costs. For commuters of $n > \bar{n}$, it is impossible to reach the bottleneck before the queue reaches the peak, which will increase the travel cost too.

2.2 User-Equilibrium Traffic Pattern in the Morning Considering Daily Travel Cost

In Sect. 2.1, we found that individual travel cost vary and depend on where you park your car. As a result, travelers don't have the same evening commute because their choice of parking spots in the morning rush hour will affected the evening commute pattern. Households are different with the evening commute in the morning rush hour, as they are free to choose when to depart and where to park.

It is assumed that each commuter takes into account his or her daily cost as a whole. In this condition, we studied a user equilibrium on morning rush hour commuting that considered the travel cost of evening commuting. Besides, in this part, our research object changes from individual commuting to household commuting. Commuters would drive to their children to school first and then go to the workplace and one household commuter have only one car.

The travel cost of commuters who leave the bottleneck at time $\vec{t}$ and with parking location n in commuting is

$$C^f\left(\vec{t}, n\right) = \alpha \frac{\vec{D}\left(\vec{t}\right)}{\vec{s}} + \lambda wn + \vec{\beta}\left(\vec{t}_1^* - \vec{t}\right) + \vec{\beta}\left(\vec{t}_2^* - \vec{t} - wn\right) + \overleftarrow{C}_n^f. \tag{7}$$

For the case of $n \le \bar{n}$, we get

$$C^f(\vec{t},n)=\alpha\frac{\vec{D}(\vec{t})}{\vec{s}}+\lambda wn+\vec{\beta}(\vec{t}_1^*-\vec{t})+\vec{\beta}(\vec{t}_2^*-\vec{t}-wn)+\frac{\alpha n}{\overleftarrow{s}}-\alpha wn+\lambda wn, \tag{8}$$

and

$$\frac{\partial C^f(\vec{t},n)}{\partial n}=\left(2\lambda-\alpha-\vec{\beta}\right)w+\frac{\alpha}{\overleftarrow{s}}>0; \tag{9}$$

For the case of $n>\bar{n}$, we can get

$$C^f(\vec{t},n)=\alpha\frac{\vec{D}(\vec{t})}{\vec{s}}+\lambda wn+\vec{\beta}(\vec{t}_1^*-\vec{t})+\vec{\beta}(\vec{t}_2^*-\vec{t}-wn)+\left(\lambda-\overleftarrow{\gamma}\right)wn+\frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}, \tag{10}$$

and

$$\frac{\partial C^f(\vec{t},n)}{\partial n}=\left(2\lambda-\vec{\beta}-\overleftarrow{\gamma}\right)w>0. \tag{11}$$

According to the assumption, household traveler goes through the bottleneck and parks as close as possible to CBD.

Firstly, we discuss the situation $n\le\bar{n}$. On account of n is depended by $p,\vec{t},\vec{t}_0^f$, we get $n(\vec{t})=p\left(\vec{t}-\vec{t}_0^f\right)$, where p means the departure rate of bottleneck and $\vec{t}_0^f$ is the initial departure time. If the bottleneck runs at/over its capacity limitation, the queue will form, then $p=\vec{s}$, otherwise $p<\overleftarrow{s}$. Commuters would make a tradeoff between the benefit of parking at a good on-street parking spot and reaching CBD close to the preferred arrival time. Once the cost saving of good on-street parking spot exceeds that of having appropriate arrival time, some household travelers will leave home earlier than others to find a convenient parking location. In this condition, no queue is formed and the leaving rate is less than the capacity of bottleneck. On the other hand, once the benefit of finding a convenient parking location is less than that of having appropriate arrival time, due to the limitation of the bottleneck capacity, the peak hours are still very crowded before the working start time, and there is a queue.

We have a presupposition, i.e. $p=\overleftarrow{s}$, and test its validity. The travel cost of household who leaves the bottleneck at time $\vec{t}$ with parking location $n(\vec{t})=p\left(\vec{t}-\vec{t}_0^f\right)$ is

$$\begin{aligned}C^f(\vec{t},n)=\alpha\frac{\vec{D}(\vec{t})}{\vec{s}}&+\left(2\lambda-\alpha-\vec{\beta}\right)w\vec{s}\left(\vec{t}-\vec{t}_0^f\right)+\vec{\beta}(\vec{t}_1^*-\vec{t})\\&+\vec{\beta}(\vec{t}_2^*-\vec{t})+\frac{\alpha}{\overleftarrow{s}}\vec{s}\left(\vec{t}-\vec{t}_0^f\right).\end{aligned} \tag{12}$$

The requirement of user-equilibrium condition is

$$\frac{dC^f(\vec{t},n)}{d\vec{t}} = \frac{\alpha}{\vec{s}}\frac{d\vec{D}(\vec{t})}{d\vec{t}} + \left(2\lambda - \alpha - \vec{\beta}\right)w\vec{s} - 2\vec{\beta} + \frac{\alpha}{\overleftarrow{s}}\vec{s} = 0. \tag{13}$$

Then we can obtain the varying rate of the queue

$$\frac{d\vec{D}(\vec{t})}{d\vec{t}} = \frac{\vec{s}}{\alpha}\left[2\vec{\beta} - \frac{\alpha}{\overleftarrow{s}}\vec{s} + \left(\alpha + \vec{\beta} - 2\lambda\right)w\vec{s}\right]. \tag{14}$$

In our study, we only analyze the usual situation. $\frac{\vec{s}}{\overleftarrow{s}} > \frac{2\vec{\beta}}{\alpha} + \left(1 + \frac{\vec{\beta}}{\alpha} - \frac{2\lambda}{\alpha}\right)w\vec{s}$ means the capacity of bottleneck in work-to-home direction isn't much greater than that in home-to-work direction. According to the assumption, we always have $\frac{d\vec{D}(\vec{t})}{d\vec{t}} < 0$, and a queue can't be negative under user-equilibrium conditions.

Hence, we can prove that the hypothesis $p = \vec{s}$ is not true and the bottleneck is operating under its capacity, thus, $p < \vec{s}$. Taking $\vec{D}(\vec{t}) = 0$ into Eq. (8), we have

$$\frac{dC^f(\vec{t},n)}{d\vec{t}} = \left[\left(2\lambda - \alpha - \vec{\beta}\right)w + \frac{\alpha}{\overleftarrow{s}}\right]\frac{dn}{d\vec{t}} - 2\vec{\beta},\ for\, n \le \bar{n}. \tag{15}$$

In user-equilibrium, $\frac{dC^f(\vec{t},n)}{d\vec{t}} = 0$, and the arrival rate is equal to the departure rate, which can be shown as follow:

$$p = \vec{r}_1^{\,f} = \frac{dn}{d\vec{t}} = \frac{2\vec{\beta}\overleftarrow{s}}{\left(2\lambda - \alpha - \vec{\beta}\right)w\overleftarrow{s} + \alpha}. \tag{16}$$

Secondly, we discuss the case of $n > \bar{n}$. Look at this situation, $n(\vec{t}) = \bar{n} + q\left(\vec{t} - \vec{t}_{\bar{n}}^{\,f}\right)$, here q is the bottleneck departure rate and $\vec{t}_{\bar{n}}^{\,f}$ is the $\bar{n}th$ commuter's departure time. If queue exists in this period, we can get $q = \vec{s}$ and the travel cost of a household is given by the following equation:

$$\begin{aligned} C^f(\vec{t},n) = {} & \alpha\frac{\vec{D}(\vec{t})}{\vec{s}} + w\left(2\lambda - \overleftarrow{\gamma}\right)\left(\bar{n} + \vec{s}\vec{t} - \vec{s}\vec{t}_{\bar{n}}^{\,f}\right) \\ & + \vec{\beta}\left[\vec{t}_1^{\,*} + \vec{t}_2^{\,*} - 2\vec{t} - w\left(\bar{n} + \vec{s}\vec{t} - \vec{s}\vec{t}_{\bar{n}}^{\,f}\right)\right] + \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}. \end{aligned} \tag{17}$$

When the system reach user-equilibrium condition, no household traveler can choose departure time to decrease the travel cost. So

$$\frac{dC^f(\vec{t},n)}{d\vec{t}} = \frac{\alpha}{\vec{s}}\frac{d\vec{D}(\vec{t})}{d\vec{t}} + \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} - 2\vec{\beta} = 0. \tag{18}$$

Then we can obtain the varying rate of the queue

$$\frac{d\vec{D}(\vec{t})}{d\vec{t}} = \frac{\vec{s}}{\alpha}\left[2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s}\right]. \tag{19}$$

The premise of a positive queue is $2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} > 0$, and we define this situation as regime f(a). In this case, the departure rate is

$$\vec{r}_2^{fa} = \frac{\alpha\vec{s}}{\alpha - 2\vec{\beta} + \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s}}. \tag{20}$$

$$TC^{fa} = N\left[\frac{\vec{\beta}}{\vec{s}} + 2\vec{\beta}w + \frac{\overleftarrow{\gamma}}{\alpha + \overleftarrow{\gamma}w\overleftarrow{s} - \alpha w\overleftarrow{s}}\left(\frac{\alpha}{\overleftarrow{s}} - \frac{\vec{\beta}}{\vec{s}} + 2\lambda w - \alpha w - \vec{\beta}w\right)\right] + \vec{\beta}\left(\vec{t}_1^* - \vec{t}_2^*\right). \tag{21}$$

In the whole system, the total social cost is

$$SC^{fa} = N^2\left[\frac{\vec{\beta}}{\vec{s}} + 2\vec{\beta}w + \frac{\overleftarrow{\gamma}}{\alpha + \overleftarrow{\gamma}w\overleftarrow{s} - \alpha w\overleftarrow{s}}\left(\frac{\alpha}{\overleftarrow{s}} - \frac{\vec{\beta}}{\vec{s}} + 2\lambda w - \alpha w - \vec{\beta}w\right)\right] + \vec{\beta}\left(\vec{t}_1^* - \vec{t}_2^*\right)N. \tag{22}$$

When $2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} \leq 0$, and we define this case as regime f(b). In this situation, there is no queue at the bottleneck, i.e. $q < \vec{s}$. Under the condition $\frac{dC^f(\vec{t},n)}{d\vec{t}} = 0$ in user-equilibrium, the arrival rate or departure rate is

$$q = \vec{r}_2^{fb} = \frac{dn}{d\vec{t}} = \frac{2\vec{\beta}}{\left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w}. \tag{23}$$

Household travel cost in regime f(b) is

$$TC^{fb} = N\left[\left(2\lambda + \vec{\beta} - \overleftarrow{\gamma}\right)w + \frac{\overleftarrow{\gamma}}{\overleftarrow{s}}\right] + \vec{\beta}\left(\vec{t}_1^* - \vec{t}_2^*\right). \tag{24}$$

In the whole system, the total social cost is

$$SC^{fb} = N^2\left[\left(2\lambda + \vec{\beta} - \overleftarrow{\gamma}\right)w + \frac{\overleftarrow{\gamma}}{\overleftarrow{s}}\right] + \vec{\beta}\left(\vec{t}_1^* - \vec{t}_2^*\right)N. \tag{25}$$

3 Morning and Evening Commuting Patterns with Time-Varying Road Tolls (r)

After the discussion of the basic situation without considered charge policy, we take the time-varying road toll into account.

(a) *Travel pattern in the evening commute with a time-varying road toll*

Under this situation, the travel cost of an individual who leaves the bottleneck at time $\overleftarrow{t}$ with a parking spot n is

$$\overleftarrow{C}^{r}\left(\overleftarrow{t}, n\right) = \lambda wn + \overleftarrow{\gamma}\left[\overleftarrow{t} - wn - \overleftarrow{t}^{*}\right] + \overleftarrow{\tau}\left(\overleftarrow{t}\right). \tag{26}$$

The last term $\overleftarrow{\tau}\left(\overleftarrow{t}\right)$ is the evening road toll according to departure time $\overleftarrow{t}$. And the efficiency loss caused by queuing in regime f is eliminated.

On the following step, we derive the condition equation which the optimal departure time must satisfy:

$$\frac{\partial \overleftarrow{C}^{r}\left(\overleftarrow{t}, n\right)}{\partial \overleftarrow{t}} = 0, \tag{27}$$

$$\frac{d\overleftarrow{\tau}\left(\overleftarrow{t}\right)}{d\overleftarrow{t}} = -\overleftarrow{\gamma}. \tag{28}$$

So, a time-varying linear toll should be charged at the bottleneck to reach the user-equilibrium. The toll can be expressed as

$$\overleftarrow{\tau}\left(\overleftarrow{t}\right) = \overleftarrow{\tau}^{0} - \overleftarrow{\gamma}\left(\overleftarrow{t} - \overleftarrow{t}^{*}\right). \tag{29}$$

According to the reality that toll can't be negative, we should set the last commuter to depart pays no toll, meanwhile the first commuter pays $\overleftarrow{\tau}^{0} = \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}$. The necessary and sufficient condition that support the arrival curve AB is $n(k) \leq \frac{k}{w\overleftarrow{s}}$, $\forall k \in [0, N]$.

Then we obtain the essential conclusion about the individual travel cost in equilibrium: $\left(\lambda - \overleftarrow{\gamma}\right)wn + \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}$.

The queue is completely eliminated by postponing the departure times of commuters because of the evening toll. The increments of total waiting time in the office is equal to the total queuing delay, and with the condition $\alpha > \overleftarrow{\gamma}$, social travel cost is lower than regime f (Fig. 3).

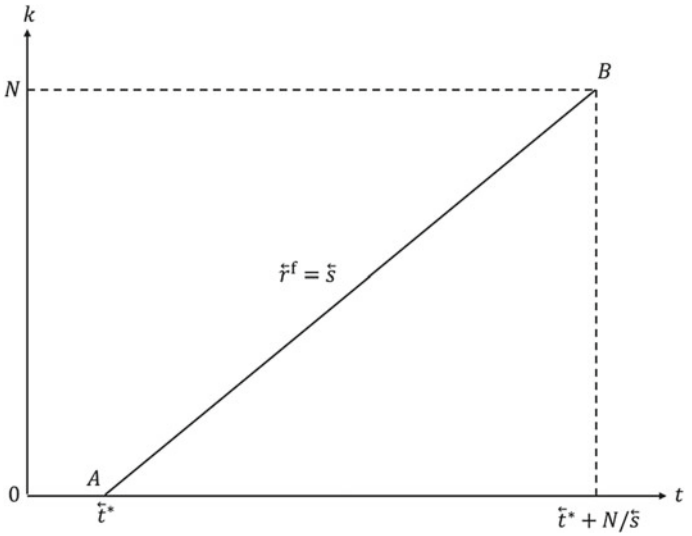


Fig. 3 User-equilibrium traffic pattern in the evening commute, regime r

(b) *Travel pattern in the morning commute with an optimal time-varying road toll*

The general formula of individual travel cost about $\vec{t}$ is

$$C^r(\vec{t}, n) = \vec{\tau}(\vec{t}) + \left(2\lambda - \overleftarrow{\gamma}\right)wn + \vec{\beta}(\vec{t}_1^* - \vec{t}) + \vec{\beta}(\vec{t}_2^* - \vec{t} - wn) + \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}. \quad (30)$$

And we have

$$\frac{\partial C^r(\vec{t}, n)}{\partial n} = \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w > 0. \quad (31)$$

Commuters would like to choose the closest park spot to the office and they pass the bottleneck in order. Hence the parking spot for the commuter who departs from the bottleneck at time $\vec{t}$ is $n(\vec{t}) = \vec{s}(\vec{t} - \vec{t}_0^r)$, here $\vec{t}_0^r$ is the earliest departure time.

So we can get the individual travel cost with parking spot n is

$$C^r(\vec{t}, n) = \vec{\tau}(\vec{t}) + \left(2\lambda - \overleftarrow{\gamma}\right)w\vec{s}(\vec{t} - \vec{t}_0^r) + \vec{\beta}(\vec{t}_1^* - \vec{t}) + \vec{\beta}\left[\vec{t}_2^* - \vec{t} - w\vec{s}(\vec{t} - \vec{t}_0^r)\right] + \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}. \quad (32)$$

In user-equilibrium condition, no one can unilaterally change departure time from home to saving cost. So

$$\frac{dC^r(\vec{t},n)}{d\vec{t}} = \frac{d\vec{\tau}(\vec{t})}{d\vec{t}} + \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} - 2\vec{\beta} = 0. \tag{33}$$

The temporally varying rate of the bottleneck tolls is given by

$$\frac{d\vec{\tau}(\vec{t})}{d\vec{t}} = 2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s}. \tag{34}$$

Further, the time dependent tolls are

$$\vec{\tau}(\vec{t}) = \vec{\tau}^0 + 2\vec{\beta}\vec{t} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s}\vec{t}, \tag{35}$$

where $\vec{\tau}^0$ means the toll of the first commuter and is non-negative. It is given by

$$\vec{\tau}^0 = \begin{cases} 0 & if\ 2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} \geq 0; \\ \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)wN - 2\vec{\beta}\frac{N}{\vec{s}} & if\ 2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} < 0. \end{cases} \tag{36}$$

In this regime, the household travel cost is

$$TC^r = \begin{cases} \vec{\beta}\left(w + \frac{2}{\vec{s}}\right)N + \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}} & if\ 2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} \geq 0; \\ \left(2\lambda - \overleftarrow{\gamma}\right)wN + \frac{N\overleftarrow{\gamma}}{\overleftarrow{s}} & if\ 2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} < 0. \end{cases} \tag{37}$$

In the whole system, the total social cost is

$$SC^r = \left[\lambda w + \frac{\overleftarrow{\gamma}}{2}\left(\frac{1}{\overleftarrow{s}} - w\right) + \frac{\vec{\beta}}{2}\left(w + \frac{2}{\vec{s}}\right)\right]N^2. \tag{38}$$

4 Time-Varying Road Tolls and Location-Dependent Parking Fees for a System Optimum (o)

In this section, we introduce location-dependent parking fees to decrease the total cost of schedule variance. We get the optimal charge policy which can achieve the optimal conditions in morning commute. Besides, the evening commute pattern is already optimized in last section.

With the parking fees, the travel cost of a household who leaves the bottleneck at time $\vec{t}$ with a parking spot n is

$$C^o(\vec{t},n)=\phi(n)+\left(2\lambda-\overleftarrow{\gamma}\right)wn+\vec{\beta}\left(\vec{t}_1^*-\vec{t}\right)+\vec{\beta}\left(\vec{t}_2^*-\vec{t}-wn\right)+\frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}. \quad (39)$$

And the reverse parking order requires

$$\frac{\partial C^o(\vec{t},n)}{\partial n}=\left(2\lambda-\vec{\beta}-\overleftarrow{\gamma}\right)w+\frac{d\phi(n)}{dn}<0. \quad (40)$$

Under the parking inversion condition, we can know the varying rate of parking fees must satisfy $\frac{d\phi(n)}{dn}<\left(\vec{\beta}+\overleftarrow{\gamma}-2\lambda\right)w$. The parking location for the household traveler who leaves the bottleneck at time $\vec{t}$ is $n(\vec{t})=N-\vec{s}(\vec{t}-\vec{t}_0^o)$, because of the difference parking order. And $\vec{t}_0^o$ is the earliest departure time, but $\frac{dn(\vec{t})}{d\vec{t}}=-\vec{s}$.

In equilibrium condition,

$$\frac{dC^o(\vec{t},n)}{d\vec{t}}=-\frac{d\phi(n)}{dn}\vec{s}-\left(2\lambda-\vec{\beta}-\overleftarrow{\gamma}\right)w\vec{s}-2\vec{\beta}=0. \quad (41)$$

The varying rate of the parking fees is given by

$$\frac{d\phi(n)}{dn}=\left(\vec{\beta}+\overleftarrow{\gamma}-2\lambda\right)w-\frac{2\vec{\beta}}{\vec{s}}. \quad (42)$$

Further, the parking fees are

$$\phi(n)=\phi^0+\left(\vec{\beta}+\overleftarrow{\gamma}-2\lambda\right)wn-\frac{2\vec{\beta}}{\vec{s}}n, \quad (43)$$

where ϕ^0 means the minimum parking fee and is non-negative. It is given by

$$\phi^0=\left(2\lambda-\vec{\beta}-\overleftarrow{\gamma}\right)wN+\frac{2\vec{\beta}}{\vec{s}}N. \quad (44)$$

In this regime, the household travel cost is shown in the following equation:

$$TC^o=N\left[\left(2\lambda-\vec{\beta}-\overleftarrow{\gamma}\right)w+\frac{2\vec{\beta}}{\vec{s}}+\frac{\overleftarrow{\gamma}}{\overleftarrow{s}}\right]. \quad (45)$$

In the whole system, the total social cost is

$$SC^o=N^2\left[\lambda w+\frac{\overleftarrow{\gamma}}{2}\left(\frac{1}{\overleftarrow{s}}-w\right)+\frac{\vec{\beta}}{2}\left(\frac{2}{\vec{s}}-w\right)\right]. \quad (46)$$

With the introduction of location-dependent parking fees, the joint strategy of road tolls and parking fees can also minimize the cost of schedule delays by changing the order of parking. They system achieves optimum.

5 Optimal Location-Dependent Parking Fees Without Roads Tolls (p)

Compared with parking fees, road pricing has disadvantage in many aspect, such as low willingness to accept, high operating cost and so on. This is why parking charges are more widely used in cities. In this section, the regime about location-dependent parking fees without road tolls in morning/evening commuting is discussed. Parking fees can only optimize the morning commute pattern which is based on the evening commuting pattern in regime f.

The travel cost of a household who leaves the bottleneck at time $\vec{t}$ with a parking spot n is

$$C^p\left(\vec{t},n\right)=\phi(n)+\lambda wn+\vec{\beta}\left(\vec{t}_1^*-\vec{t}\right)+\vec{\beta}\left(\vec{t}_2^*-\vec{t}-wn\right)+\overleftarrow{C}_n^f. \tag{47}$$

In the case of $n\le\bar{n}$, we have

$$C^p\left(\vec{t},n\right)=\phi(n)+\lambda wn+\vec{\beta}\left(\vec{t}_1^*-\vec{t}\right)+\vec{\beta}\left(\vec{t}_2^*-\vec{t}-wn\right)+\frac{\alpha n}{\overleftarrow{s}}-\alpha wn+\lambda wn, \tag{48}$$

and

$$\frac{\partial C^p\left(\vec{t},n\right)}{\partial n}=\frac{d\phi(n)}{dn}+\left(2\lambda-\alpha-\vec{\beta}\right)w+\frac{\alpha}{\overleftarrow{s}}; \tag{49}$$

and in the case of $n>\bar{n}$, we know

$$C^p\left(\vec{t},n\right)=\phi(n)+\left(2\lambda-\overleftarrow{\gamma}\right)wn+\vec{\beta}\left(\vec{t}_1^*-\vec{t}\right)+\vec{\beta}\left(\vec{t}_2^*-\vec{t}-wn\right)+\frac{N\overleftarrow{\gamma}}{\overleftarrow{s}}, \tag{50}$$

and

$$\frac{\partial C^p\left(\vec{t},n\right)}{\partial n}=\frac{d\phi(n)}{dn}+\left(2\lambda-\vec{\beta}-\overleftarrow{\gamma}\right)w. \tag{51}$$

The reverse parking order requires $\frac{\partial C^p\left(\vec{t},n\right)}{\partial n}<0$, therefore the varying rate of parking fees must satisfy the following conditions:

$$\frac{d\phi(n)}{dn} < \begin{cases} \left(\alpha + \vec{\beta} - 2\lambda\right)w - \frac{\alpha}{\overleftarrow{s}} & if\ n \le \bar{n}; \\ \left(\vec{\beta} + \overleftarrow{\gamma} - 2\lambda\right)w & if\ n > \bar{n}. \end{cases} \tag{52}$$

Since the parking order is reversed and commuters reach the parking spot in the rate of $\vec{s}$, the parking spot for the traveler who departs from the bottleneck at time $\vec{t}$ is $n(\vec{t}) = N - \vec{s}(\vec{t} - \vec{t}_0^p)$, where $\vec{t}_0^p$ is the earliest leaving time. And so $\frac{dn(\vec{t})}{d\vec{t}} = -\vec{s}$.

In the case of $n \le \bar{n}$, user-equilibrium condition has to satisfy

$$\frac{dC^p(\vec{t}, n)}{d\vec{t}} = -\frac{d\phi(n)}{dn}\vec{s} - \left(2\lambda - \alpha - \vec{\beta}\right)w\vec{s} - 2\vec{\beta} - \frac{\alpha}{\overleftarrow{s}}\vec{s} = 0. \tag{53}$$

The varying rate of the parking fees is given by

$$\frac{d\phi(n)}{dn} = \left(\alpha + \vec{\beta} - 2\lambda\right)w - \frac{2\vec{\beta}}{\vec{s}} - \frac{\alpha}{\overleftarrow{s}}. \tag{54}$$

In the case of $n > \bar{n}$, user-equilibrium condition has to satisfy

$$\frac{dC^p(\vec{t}, n)}{d\vec{t}} = -\frac{d\phi(n)}{dn}\vec{s} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} - 2\vec{\beta} = 0. \tag{55}$$

The varying rate of the parking fees is given by

$$\frac{d\phi(n)}{dn} = \left(\vec{\beta} + \overleftarrow{\gamma} - 2\lambda\right)w - \frac{2\vec{\beta}}{\vec{s}}. \tag{56}$$

In user-equilibrium, the varying rate of parking fees satisfies condition (52). With the assumption that the first commuter don't have to pay, then we obtain the minimal non-negative parking fees for all locations.

$$\phi(n) = \begin{cases} \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w(N - \bar{n}) + \frac{2\vec{\beta}(N-\bar{n})}{\vec{s}} + \left[\left(2\lambda - \alpha - \vec{\beta}\right) + \frac{\alpha}{\overleftarrow{s}}\right](\bar{n} - n) & if\ \mathrm{n} > \bar{\mathrm{n}}; \\ \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w(N - n) + \frac{2\vec{\beta}(N-n)}{\vec{s}} & if\ \mathrm{n} > \bar{\mathrm{n}}. \end{cases} \tag{57}$$

The travel cost of household commuter is shown in the following equation:

$$TC^p = N\left[\left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w + \frac{2\vec{\beta}}{\vec{s}} + \frac{\overleftarrow{\gamma}}{\overleftarrow{s}}\right]. \tag{58}$$

In the whole system, the total social cost is

$$SC^p = \frac{N^2}{2}\left[\left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w + \frac{2\vec{\beta}}{\vec{s}} + \frac{\overleftarrow{\gamma}}{\overleftarrow{s}}\left(2 - \frac{\overleftarrow{\gamma}}{\alpha + \overleftarrow{\gamma}w\overleftarrow{s} - \alpha w\overleftarrow{s}}\right)\right]. \quad (59)$$

6 Numerical Examples

In this section, we use a numerical example to illustrate the research results. We assume there is a bottleneck before both trip destinations of household travels in the morning, i.e., school and workplace. Suppose a fixed number of $N = 1.0 \times 10^4$ household commuters need to depart from home to their children's school firstly and then to their workplaces every morning and return home in the evening. The capacity of the bottleneck is $\vec{s} = \overleftarrow{s} = 1.0 \times 10^2$ veh/min. The walking time for passing one parking point is $w = 0.001$ min and value of unit walking time $\lambda = 2.0\$/\text{min}$. We set the cost of unit travel time $\alpha = 0.6\$/\text{min}$, the penalty cost per unit time for early arrival is $\vec{\beta} = 0.3\$/\text{min}$ and the cost per unit time for leaving late is $\overleftarrow{\gamma} = 0.3\$/\text{min}$.

The school start time is $\vec{t}_1^* = 8{:}30$ am, official beginning time is $\vec{t}_2^* = 9{:}$ am and the work ending time is $\overleftarrow{t}^* = 17{:}00$ pm (Table 1).

According to the numerical setting, we can know that $2\vec{\beta} - \left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w\vec{s} < 0$. So we use equation SC^{fb} when it comes to calculating the social costs under regime f (Table 2).

The total social cost of regime f is the highest, and the introduction of different charging policies can reduce the total social cost. The regime with road tolls policy can significantly reduce household and social cost. Through the example, we can

Table 1 Total social costs for various regimes

Regimes	Social cost
No road tolls or parking fees	$SC^{fa} = N^2\left[\frac{\vec{\beta}}{\vec{s}} + 2\vec{\beta}w + \frac{\overleftarrow{\gamma}}{\alpha + \overleftarrow{\gamma}w\overleftarrow{s} - \alpha w\overleftarrow{s}}\left(\frac{\alpha}{\overleftarrow{s}} - \frac{\vec{\beta}}{\vec{s}} + 2\lambda w - \alpha w - \vec{\beta}w\right)\right] + \vec{\beta}(\vec{t}_1^* - \vec{t}_2^*)N$
	$SC^{fb} = N^2\left[\left(2\lambda + \vec{\beta} - \overleftarrow{\gamma}\right)w + \frac{\overleftarrow{\gamma}}{\overleftarrow{s}}\right] + \vec{\beta}(\vec{t}_1^* - \vec{t}_2^*)N$
Optimal road tolls	$SC^r = \left[\lambda w + \frac{\overleftarrow{\gamma}}{2}\left(\frac{1}{\overleftarrow{s}} - w\right) + \frac{\vec{\beta}}{2}\left(w + \frac{2}{\vec{s}}\right)\right]N^2$
Optimal road tolls and parking fees	$SC^o = N^2\left[\lambda w + \frac{\overleftarrow{\gamma}}{2}\left(\frac{1}{\overleftarrow{s}} - w\right) + \frac{\vec{\beta}}{2}\left(\frac{2}{\vec{s}} - w\right)\right]$
Optimal parking fees without roads tolls	$SC^p = \frac{N^2}{2}\left[\left(2\lambda - \vec{\beta} - \overleftarrow{\gamma}\right)w + \frac{2\vec{\beta}}{\vec{s}} + \frac{\overleftarrow{\gamma}}{\overleftarrow{s}}\left(2 - \frac{\overleftarrow{\gamma}}{\alpha + \overleftarrow{\gamma}w\overleftarrow{s} - \alpha w\overleftarrow{s}}\right)\right]$

Table 2 Numerical results

Regime	Household cost ($)	Social cost ($)
f	79	790,000
r	67	650,000
o	124	620,000
p	124	691,053

prove that regime o is optimal from the perspective of total social cost. This regime reduces the cost of waiting time and schedule delay.

Compared to regime o, regime p has the same household travel cost but the higher social cost. If road tolls are not credible in some cities, such a policy can also greatly reduce the overall social cost.

7 Conclusions

In this paper, a joint morning and evening rush-hour commute for household is investigated. We assumed that the daily commute link is connected by home, school and workplace. Firstly, we derived the evening work-to-home commuting pattern for household, by analyzing the travel costs with different parking locations. Then the morning home-to-work commuting pattern was developed through the system total travel cost in a whole day. Furthermore, we design pricing schemes, i.e., various road toll and/or parking fee, to improve the system efficiency. It is found that the road tolls can eliminate the queue behind the bottleneck and even reduce the penalty. Moreover, the joint strategy of road tolls and parking fees can also minimize the cost of schedule delays by changing the order of parking. Consider the fact that the road tolls is difficult to implement, a location-dependent parking fees was proposed, which can optimize the morning commute pattern by eliminating the queue in morning travel and minimize the schedule delay cost with reverse order of parking.

Furthermore, we found that the road tolls on morning can only affect morning traffic patterns, but the road tolls on evening can affect both morning and evening traffic patterns. Compared with road tolls, the location-dependent parking fees are more effective. Since the location-dependent parking fee can minimize the cost of schedule delays and the total travel cost by changing the order of parking We also found that the location-dependent parking fees only affects the morning commute but has no influence on the evening commute. It is found that the optimal daily travel can be achieved by adopting the joint parking fees and road tolls scheme.

Further research may introduce user heterogeneity based on different activity. For example, compared with work commuters, household travelers need complete two destination, i.e., school and workplace. It is meaningful to consider both work commute and household commute into the analysis, and distinguish the traffic pattern for traffic congestion. Besides, it is also of interest to consider modal split, demand uncertainty and flexible work schedules.

Acknowledgements The research described in this paper was jointly supported by grants from the National Natural Science Foundation of China (71501012), the Beijing Social Science Foundation (16GLC054).

References

1. Vickrey WS (1969) Congestion theory and transportation investment. Am Econ Rev 59:251–261
2. Arnott R, De Palma A, Lindsey R (1990) Economics of a bottleneck. J Urban Econ 27(1):111–130
3. Arnott R, de Palma A, Lindsey R (1998) Recent developments in the bottleneck model. In: Button KJ, Verhoef ET (eds) Road pricing, traffic congestion and the environment: issues of efficiency and social feasibility. Edward Elga, Aldershot, pp 161–179
4. de Palma A, Fosgerau M (2011) Dynamic traffic modeling. In: de Palma A, Lindsey R, Quinet E, Vickeman R (eds) Handbook in transport economics. Edvard Elgar, Cheltenham, UK, pp 29–37
5. Lindsey R, van den Berg V, Verhoef ET (2012) Step tolling with bottleneck queuing congestion. J Urban Econ 72(1):46–59
6. Xiao LL, Liu TL, Huang HJ (2019) Tradable permit schemes for managing morning commute with carpool under parking space constraint. Transportation. https://doi.org/10.1007/s11116-019-09982-w
7. Arnott R, De Palma A, Lindsey R (1991) A temporal and spatial equilibrium analysis of commuter parking. J Public Econ 45:301–335
8. Zhang X, Huang HJ, Zhang HM (2008) Integrated daily commuting patterns and optimal road tolls and parking fees in a linear city. Transp Res Part B 42:38–56
9. Jia Z, Wang DZW, Cai X (2016) Traffic managements for household travels in congested morning commute. Transp Res Part E Logistics Transp Rev 91:173–189
10. Liu W, Zhang F, Yang H (2016) Modeling and managing morning commute with both household and individual travels [J]. Transport Res B: Meth

Airfield Smart Operations Management and Application of Shared Services

Zhang Rui

Abstract Nowadays, traditional ground-service model cannot meet the requirements of multiple-runway and multiple-terminal operations. To solve the problem, we introduce new technology and applications to change the traditional safety operations management in all aspects and multiple perspectives. This article takes unpowered equipment management and smart transportation as examples to make a research on shared services which based on Airfield Smart Operations Management.

Keywords Unpowered equipment · Smart airport · Intelligent management · Transportation service

Airfield is the area for safety take-off and landing, including runway, runway strip, RESA, stop way, clearway, taxiway, apron, drainage system, soil surface area, security system, inspection tour road and any other system. Apron is the most important area of airfield, which provides aircraft maintenance and ground-service. Different management unit work together to provide service for each aircraft, such as airport company, air traffic control, airlines, aviation oil company, ground-service, aviation logistics and so on. These units may belong to the same company but different aircraft service procedures or same procedures but different companies. No matter what patterns, all the unites use general aviation equipment and comply with same procedures. Based on just mentioned and shared concept, we try to break boundaries and limit to increase the utilization of airfield idle resources by applying of advanced management methods and technologies. At the same time, we will use Internet Plus strategy and Big Data to make a further research on resources utilization for future development. We expect to share and integrate resource to have a breakthrough and construct an airport of safety, environmentally friendly, intelligence and humanity. The author takes unpowered equipment management and smart transportation as examples to have a brief discussion about the application of shared concept in airfield smart management.

Z. Rui (✉)
Civil Aviation Administration of China, Beijing 100710, China
e-mail: tukun88811@126.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_105

1 Intelligence Sharing Platform of Unpowered Equipment

The operation of unpowered equipment occupies a considerable proportion in the ground support of flights. At present, there are at least 6000 unpowered equipments at airports above 4F class in China. A large number of unpowered equipment groups involve a series of problems such as investment, daily maintenance, and parking resources. Taking Kunming Airport as an example, the configuration method of the unpowered equipment on the apron is configured according to each airline support area and long-term human experience. Without scientific calculation, it is easy to cause uneven allocation of resources. Because of the large number of equipments, the space on the apron is occupied and the resources are congested and wasted. In other condition, equipments are in such a shortage to transferred in time and affects ground support of flights. In addition, many units are lagging behind in daily maintenance management after the equipment is put into use, and the most common is that the equipments are placed in areas not permitted. For example, the staff did not return the equipments to the designated areas after using, which caused the disorder. Some staffs do not lock the support feet, place wheel stop and install chain as required. It will lead to the risk of equipments being blown away by strong winds (Fig. 1). Secondly, airlines tend to maximize their spending on equipments due to the requirement of the type and quantity of equipment for aircraft support, and excessive equipments carry a huge cost of capital. These are all wastes caused by the lack of sharing concept. Finally, the equipment company does not have unified management standards, nor develop a reliable maintenance program, resulting in the equipment failure rate increasing year by year, and there are many old and damaged equipments are still working with high security risks.

Similar problems exist in the management of unpowered equipment. From the view of equipment property rights, there are differences in manufacturing standards of various types. If the equipment has to serve different types of aircraft, purchasing a complete set of equipment will result in higher procurement costs, and daily maintenance and management also require a certain amount of human costs. From the view of equipment users, the management confusion caused by random placement will lead to lower employee satisfaction with the use of equipment. Long-distance dragging, equipment malfunction occur frequently. From the perspective of equipment management department, most of the unpowered equipment is placed in an open area, although they have property rights, but there are also cases of departments

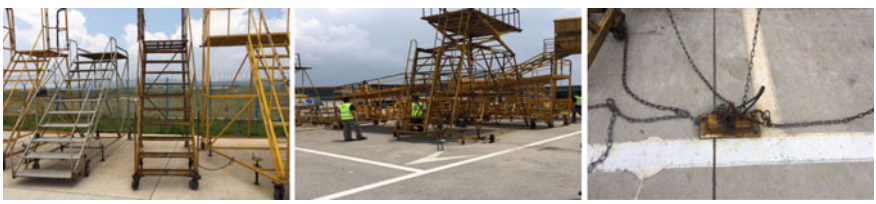


Fig. 1 Placement of the unpowered equipment

borrowing from each other, which will lead to problems such as unsafe return or difficulty in finding the directly responsible person after an accident. In the wild apron, even if the company's equipment is put to the best of its ability, there will not be satisfactory, but occupies a lot of ground resources and human resources of the airport. From the view of equipment manufacturers, the maintenance of equipment cannot be effectively guaranteed because of the lack of uniform manufacturing standards.

Therefore, how to scientifically manage the unpowered equipment in the airfield and reduce the safety hazards has become an urgent problem to be solved at this stage. As the status of idle resources of the same type, the author believes that we should imply sharing concept into the whole management of unpowered equipment, which means unified the introduction, launch and routine maintenance of unpowered equipment management by one party. All units that need to use such equipment can share usage rights through cooperative consumption or collaborative consumption. The management party relies on the fixed end of the device and the mobile device, and utilizes network technology such as device resource visualization, GPS, and user experience evaluation system to integrate idle devices in the apron, thereby achieving optimal resources for "best use" and "distribution on demand", and according to the area, unit, time, frequency, etc. of the equipment used to analyze the entire apron operations in terms of unpowered equipment protection. Through security control, device sharing and "Internet + Big Data", the following goals are achieved:

1. Real-time monitoring of using unpowered equipment in the airfield by intelligent management.
2. Using big data to analyze and decide the placement and configuration of the equipment in the airfield.
3. Implementing management regulations and norms into the daily work by informationize management.
4. To manage the whole life cycle of unpowered equipment in airfield by using electronic identity tag and intelligent management.
5. Secondly, the key points and modes of management need to be sorted out according to the equipment and personnel. In terms of the equipment management, a full life cycle management for all unpowered equipment in the airfield should be established, including equipment manufacturing, admission, utilization and scrap. The usage of equipment will be improved by combining big data analysis. In terms of the personnel management, it mainly focuses on the personnel authority and operation norms. In addition, in view of the standard formulation, it is necessary to realize the unification of the standards, including equipment manufacturing process, equipment fixing, placing, equipment use, maintenance and patrol. This part of the work can be achieved by building an intelligent platform, which composed of unpowered equipment, management platform and mobile terminal (eg. Figs. 2 and 3).

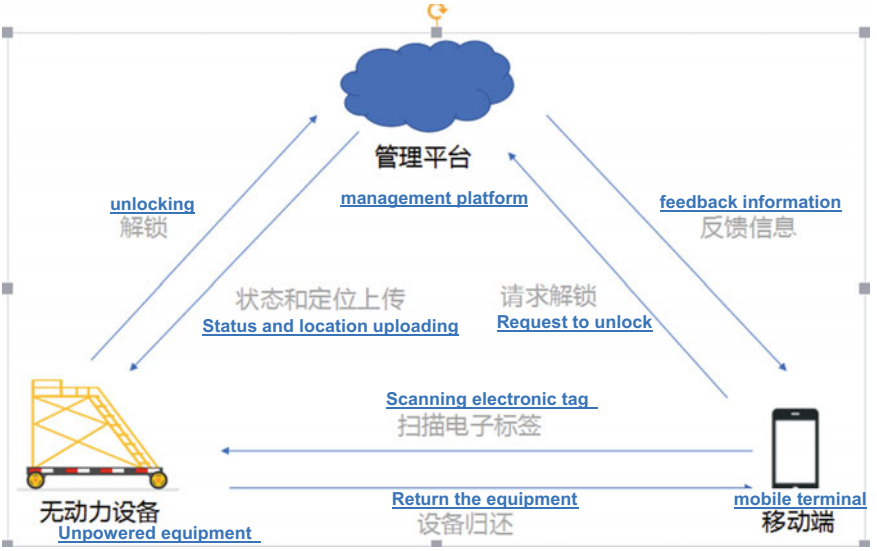


Fig. 2 The vision of intelligent management platform for non-power equipment

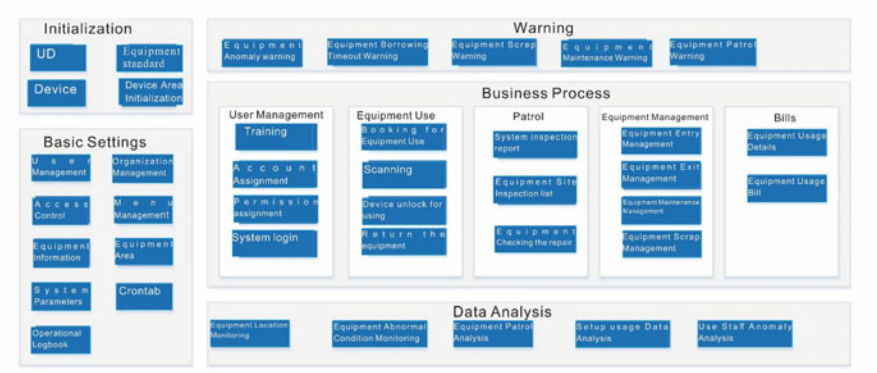


Fig. 3 The database structure

After setting up the intelligent management platform, it will concentrate on achieving the following objectives: First, to achieve accurate positioning. Managing people and equipment through technical means, detect and timely report the abnormal equipment to the equipment unit; Secondly, to standardize area. In order to reduce the potential safety hazards in the field, apron resources should be rationally used and analyzed by using the big data to minimizing the number of equipment on standby and improving the efficiency of utilization. Thirdly, to reduce costs. Reduce the management cost by scientific methods, which are provided by the big data; reduce the cost

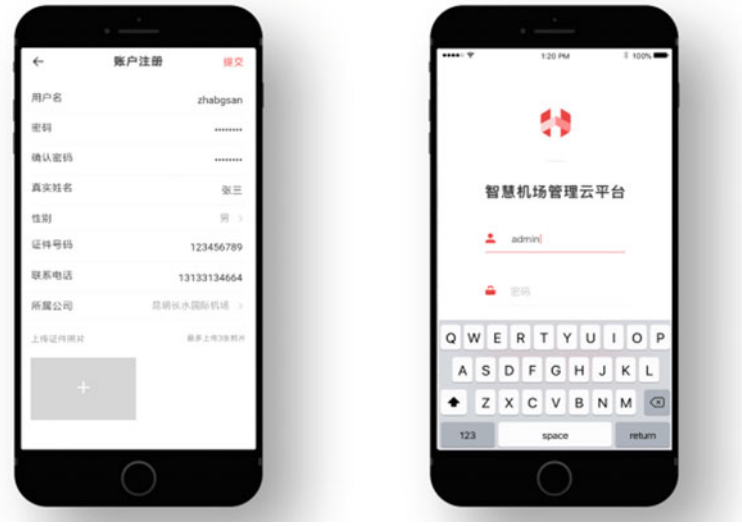


Fig. 4 Mobile smart airport management app system

of equipment manufacturing by formulating unified processes; reduce the use cost of equipment through strict management system. Fourth, equipment manufacturing standardization, i.e. using big data analysis to obtain the equipment maintenance, scrap cycle and monitoring the complete life cycle of equipment from manufacturing to scrap, through data analysis, calculate equipment manufacturing standards, meanwhile establish uniform manufacturing standards for non-powered equipment (Figs. 4 and 5).

2 Airfield Smart Vehicle Operations

Through the analysis of intelligent management of non-power equipment, it is easy to see the duplication of the diver's airfield support systems, which result in the waste of resources; on the other hand, due to the different operation department has the different peak hours, the resources may face a shortage, and the quality of operation may be debased. Therefore, the smart operation management in airfield is not only an application of informationize, but also an integration of "Internet + Big Data" and airport's Operations, and its essence is innovation and sharing.

As mentioned above, sharing is one important component of the smart airfield operations management. The larger the platform is, the more people will be affected, and the more energy will be showed up. If the main users of non-power equipment

Fig. 5 Equipment borrowing

come from maintenance units, then the smart transportation in the airfield can integrate all the staff into one group, it is an indispensable and important object in the construction of intelligent airport, and is also a concentrated reflection for the sharing economy in the airfield. Referring to the reality of Kunming Changshui International Airport, the author takes "smart transportation" as a research object, and a discussion of sharing economy in the airfield.

3 Travel Analysis for the Passengers in Airfield

Because the flight area is relatively independent and closed, the activities of the front-line working staff, crew members and passengers are limited. According to the current operation status of Kunming Airport, there are different travel difficulties and needs for these three groups of people.

3.1 Working Staff Transportation

The aircraft's station-crossing operations are concentrated on one parking apron. Since numerous personnel and facilities work within a limited time and range, spacious apron is turned into a very complex and limited area. In addition, the aircraft's operational support involves multiple departments, such as airports, airlines, jet fuel, aviation food, aircraft maintenance, ground service, logistics, a variety of support vehicles and related personnel. The attribution of personnel and vehicles results in

more complicated operation process and communication links and largely restricted travel and support operations of the staff in the apron. At present, the number of working staff in the flight area of Kunming Airport is about 13,000 (based on the scope of the flight zone in the quarantine zone) and does not include the crew members of each airline. At the same time, it is showed from the data that the number of working staff who enters the flight area is about 6000 per day, and the number of crew member is about 2400. In the entire Kunming Airport, 1096 vehicles are with Grade-D civil aviation license. The vehicle in the control zone has a limited time pass for 573. There are 42 temporary vehicle passes. In other words, 1711 vehicles can travel in the flight zone totally. The vehicles that can be used for ferry employees are about 70% except for special vehicles. At the same time, they must have other business support functions. It is inferred that there is a big contradiction between the number and scope of personnel in the flight area and the number of self-provided vehicles.

The personnel and company may face the fact that the vehicle cannot be used in a certain period of time so that the needs of both parties are affected. In addition, because of dispersed working staff, different destinations, it is necessary to ferry for a long time. As a result, it occupies the vehicle resources and increases the traffic pressure on the apron service lane, which increase the occurrence of incident.

3.2 Crew Members Transportation

The vehicle support for crew members is one of the significant conditions for normal operation of the flight. Its service and efficiency are also the user experience emphasized by crew members of each airline. Thus, the various airlines have put heavy investment in manpower and material resources in order to ensure the vehicle for crew members. However, the vehicle utilization rate for crew members of most of the airlines is relatively low. Other problems also appear such as unsatisfactory pick-up and delivery, high probability of empty cars before and after the task, and untimely deployment, wasteful resources, low efficiency, lagging service levels.

3.3 Passengers (Domestic Passengers) Transportation

The high-end customized service from the apron to the isolation area is still in a blank for the domestic airport service. However, as the society develops, consumers' demand for service quality and level has been constantly increasing. It becomes a trend for customers to use a reasonable price to increase convenience and comfort for travel. It is expected that after the project is finally promoted, the market and benefit space will be great.

4 Smart and Green Transportation Plan

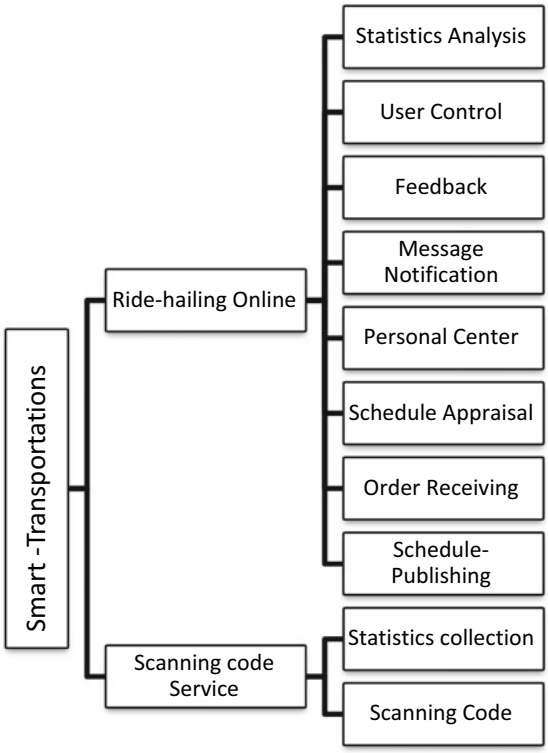
Smart and Green Transportation Plan of Kunming Airport (hereinafter referred to as Transportation service) was started in February 2018 for solving the transportation problem of apron staff at the beginning. Based on the idea, now the plan intends to provide better service for staff, crew members and passengers from 5 considerations as energy saving, resource management, pressure reducing, market development and happiness increasing.

Transportation service plan aims to solve the conflict between flight operation and transportation for staff and crew members in airfield. Considering resources management, one is to provide more means of transportation to change the way that each unite has its own transportation. Second is to ensure more vehicles can be used for flight guarantee to increase the operation efficiency. Third is to provide public transportation to reduce the frequency use of vehicles for picking up staff. Fourth is to use new energy shuttle bus for a large-scale energy saving in airfield. Fifth is to create new means of transportation to realize the service provided from inside apron to outside apron.

The first stage of the project (problem of staff transportation) has been completed on software designing and service scope setting. The second stage (problem of crew member transportation) and the third stage (problem of passenger transportation) are implemented as planned after the conditions of each previous stage are mature.

4.1 Service Plan for Staff

At present, solution of staff transportation has been on trial operation stage in Kunming Airport since October 1, 2018. The service plan on this stage includes two functions, ‘Ride-hailing Online’ and ‘Scanning Code for Service’. Staff can take the free shuttles on peak hours through scanning code on the ‘Changshui Changzhun’ APP and they can also make a car reservation on App on off-peak hours, which similar to ‘Didi’ APP. These two functions can basically satisfy the transportation demands of staff in airfield.



(System service module)

(Advertisement for trial operation)

(Staff in shuttle bus)

4.2 Crew Member Service Plan

Depending on smart and green transportation service, the airport plans to entrust a third-party company to provide exclusive transportation service, including pre-reservation or fixed transportation for crew member. For having guarantee capability and resource advantages, airport try to optimize the allocation of resources to reduce the amount of ground-service unit and to minimize the operation risk of airlines in airfield in order to advance guarantee capability and service quality. Meanwhile, the plan significantly reduces the requirement of vehicles, drivers and other thing to realize the 'intensive management' of airport, airlines and ground service companies, which is easy to promote and worth further promotion.

4.3 Passenger Service Plan

Focus on more convenient way of travel for domestic passengers, smart and green transportation system may allow passengers to ride-hailing online. Having received the reservation messages, staff arranges driver and vehicle to pick up passengers off the plane, which largely shorten the arrival time and conveniently connect to taxi or bus. So passengers can enjoy better service provided by airport.

The second phase (Airline Crew Airfield Transportation Services) and the third phase (Passenger Airfield Shuttle Services) will implement step by step when every previous stage was fulfilled. Such as when the two Ground Services companies have equipped with sufficient resources, the online car-hailing services are efficient enough and the passengers are satisfied these services.

Airfield transportation services for airport employees, airline crews and passengers seems to be a simple transportation matters. However, it is a significant consideration for building up an intelligent airport according to the five factors of safety risk, operational efficiency, cost, employee benefits and resource utilization. The essence of airfield transportation services is to provide more choices for our employees and not only improve the quality of transportation in the airfield area but also the airport transportation services in the whole and may achieve better results in the short term. Such as increase the efficiency of the apron operations, conserve energy and reduce emission to achieve low-carbon economy.

Kunming Airport has been ahead ideologically and launched the trial operation of Smart And Green Transportation Services on October 1st by relying on shared economy and introducing third-party forces. The author believes that promote new things needs time. The airport can adjust and improve the smart transportation system according to the data collected during the trial operation. Putting this system into trial operation, the employees can experience the convenience and efficiency of shared economy, which is helpful for the practical application in the near future.

Through the above analysis, it is easy to see that both the airline crew and the airport employees have great demand for smart transportation which is still a blank

in the industry in China at this moment. Whoever can take the lead in exploration in this field will seize the opportunity in the era of innovation. Smart transportation in the airfield area embodies the idea of smart sharing, and it is also an effective way to practice green development such as implementing energy saving and consumption reduction measures and the use of electric vehicles and clean energy charging piles.

At present, domestic airports have adopted the multi-terminal, multi-runway operations mode. With the rapid development of civil aviation in China, three runways, four runways and even more runways will become normal. The scale of the airfield area is expanding, the cost will increase, the difficulty of management will also increase, the single operations grantee body will become weaker and weaker, the concept of co-negotiation, co-construction, sharing and win-win will become the purpose of the operations cooperation of the airfield. The concept of shared economy will surely be put into practice. After the trial operation is completed, the service can also be implemented to the management of ferry buses, snow removal operations etc. Through the establishment of a resource sharing and scheduling platform in the airport and with several Ground Services companies participate in this project, the quality and efficiency of the services can be guaranteed, and the maximization of ground resources can be realized at the same time.

We should foster new momentum with information technology, promote new development with new momentum, and create new brilliance with new development. High-quality development of civil aviation is not empty talk, it is essentially inseparable from advanced concepts and means. Based on the intelligent operations and management of the airfield, the idea of building a shared economy in the airfield will become an important foundation and grasp of building a strong civil aviation country, and ultimately realize the transformation and upgrading of civil aviation with the help of science and technology, and accelerate the strategic landing of the dream of a strong civil aviation country.

Evaluating the Impact of Traffic Congestion on Mid-block Fine Particulate Matter Concentrations on an Urban Arterial

Xiaonian Shan, Changjiang Zheng and Xiaoli Zhang

Abstract The primary objective of this paper is to evaluate the impact of traffic congestion on mid-block fine particulate matter ($PM_{2.5}$) concentrations on an urban arterial. Data of mid-block and background $PM_{2.5}$ concentrations were collected second by second during peak and non-peak hours on an urban arterial. Then micro traffic conditions were extracted from videos at ten seconds intervals, including traffic volume, traffic flow speed and high-duty vehicle fraction. Results showed that traffic volume had significant influence on mid-block $PM_{2.5}$ concentrations. Mid-block $PM_{2.5}$ concentrations were not correlated with traffic level of service. Furthermore, a modified passenger car equivalent was calculated from the aspect of contribution on $PM_{2.5}$ concentrations using multiple linear regressions model. Then a comprehensive model was established to model the impact of micro traffic conditions on $PM_{2.5}$ concentrations. Results of the comprehensive model showed that $PM_{2.5}$ concentrations increased with the increase of total volume or heavy-duty vehicle fraction. Besides, low traffic flow speed resulted in high $PM_{2.5}$ emission factor, leading to the increase of $PM_{2.5}$ concentrations. The findings of this study can help better understand traffic congestion and micro traffic conditions on $PM_{2.5}$ concentrations.

Keywords Traffic congestion · $PM_{2.5}$ concentrations · Micro traffic conditions · Multiple linear regressions model

1 Introduction

There is evidence that fine particulate matter (particles with aerodynamic diameter less than 2.5 um, $PM_{2.5}$) is harmful for humans which can penetrate the respiratory tract and have toxic effects in heart and lungs [1]. Therefore, many countries have a $PM_{2.5}$ concentration standard for suitable living urban areas. In China, the value

X. Shan (✉) · C. Zheng · X. Zhang
College of Civil and Transportation Engineering, Hohai University,
Nanjing 210098, People's Republic of China
e-mail: shanxiaonian@hhu.edu.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_106

of average 24-h $PM_{2.5}$ concentration is 75 ug/m^3 established in 2012 [2]; while the value in United States is only 35 ug/m^3 proposed in 2012 [3].

Meanwhile, vehicle emission has been recognized as one of major contributors to $PM_{2.5}$ pollution in urban areas. For example, vehicle emission occupies 31.3% of $PM_{2.5}$ in Beijing [4]; the mobile source contributes 29.2% of $PM_{2.5}$ in Shanghai [5] and vehicle emission contributes 24.6% of $PM_{2.5}$ in Nanjing [6]. Evaluating the impact of traffic features to $PM_{2.5}$ concentration helps better understand what and how traffic factors contribute to the air pollution and related public health problem.

Previously, numerous studies regarding vehicle emission contributing to $PM_{2.5}$ concentration have been conducted. Previous studies mainly focused on two aspects to illustrate the influence of vehicle on $PM_{2.5}$ concentration: (1) forecasting $PM_{2.5}$ concentration with considering traffic volume directly; and (2) estimating the $PM_{2.5}$ concentration with combining the vehicle emission model and dispersion model. In the former subject, $PM_{2.5}$ (or PM_{10}) concentration was usually forecasted based on the former concentration and the current traffic volume [7–11]. The difference to different researches reflects in the length of time to predict $PM_{2.5}$ (or PM_{10}) concentration, such as 10-s $PM_{2.5}$ [7], 5-min $PM_{2.5}$ [8], 1-h PM_{10} [9], and 24-h PM_{10} [10]. Especially, Moore et al. analyzed the effects of traffic volume to $PM_{2.5}$ concentration at 10 s intervals and established a statistical model using ordinary least regression without considering the traffic flow speed [7]. He et al. predicted the PM_{10} concentration with considering the effect of instantaneous velocity and acceleration [11].

Some other researchers estimated the air pollutant concentrations with combining the vehicle emission model and dispersion model [12–17]. Vehicle emission models were established considering the different vehicle conditions, such as Motor Vehicle Emission Simulator (MOVES), Comprehensive Modal Emission Model (CMEM) and so on. Liu et al. estimated the NO_2 concentrations using MOVES and air dispersion model based on taxi GPS data [12]. Obviously, the estimated concentration cannot fit the field observation very well because of the complex dispersion model and traffic conditions. Su et al. established an integrated approach to estimate the pedestrian exposure to roadside with combining the traffic simulation software, emission model and dispersion model [13]. Chai et al. evaluated the impact of traffic volume and speed on carbon monoxide dispersion [14]. Some people also explored the veracity of emission factors using MOVES based on field measurements [15, 16].

In previous observational studies, the relationship between traffic congestion and $PM_{2.5}$ concentration is not clear. To micro traffic condition, only traffic volume was extracted to present the vehicle influence on $PM_{2.5}$ concentration because of data processing and study's time duration. However, vehicle emission factors are different to different types of vehicle, and traveling speed also has significant impact on vehicle emission factors. Furthermore, the technology chain is too long from traffic condition to $PM_{2.5}$ concentration resulting in poor fitting results.

The primary objective of this study is to evaluate the impact of traffic congestion on mid-block $PM_{2.5}$ concentrations based on field investigation. More specifically, this study includes the following two tasks: (1) to analyze the effect of traffic congestion on $PM_{2.5}$ concentrations; and (2) to establish a comprehensive model to evaluate

the effects of micro traffic conditions on $PM_{2.5}$ concentrations. In the following section, data collection and data preparation are introduced. In Sect. 3, the relationship between traffic congestion and $PM_{2.5}$ concentrations is analyzed. Section 4 models the effect of micro traffic conditions on $PM_{2.5}$ concentrations. The paper ends with brief concluding remarks in Sect. 5.

2 Data Collection and Preparation

2.1 Data Investigation Site and Data Collection

In this study, Cao'an Road, an urban arterial in Shanghai with 7 bidirectional lanes is selected as the field observation site, as shown in Fig. 1. The mid-block monitor was located in the center median divider to eliminate the influence of bidirectional traffic flow. The background monitor was located at the around residential area without the influence of vehicles. And the background monitor location was with 70 m to Cao'an Road and 120 m to Jia'song Road. Furthermore, Vehicle lanes are

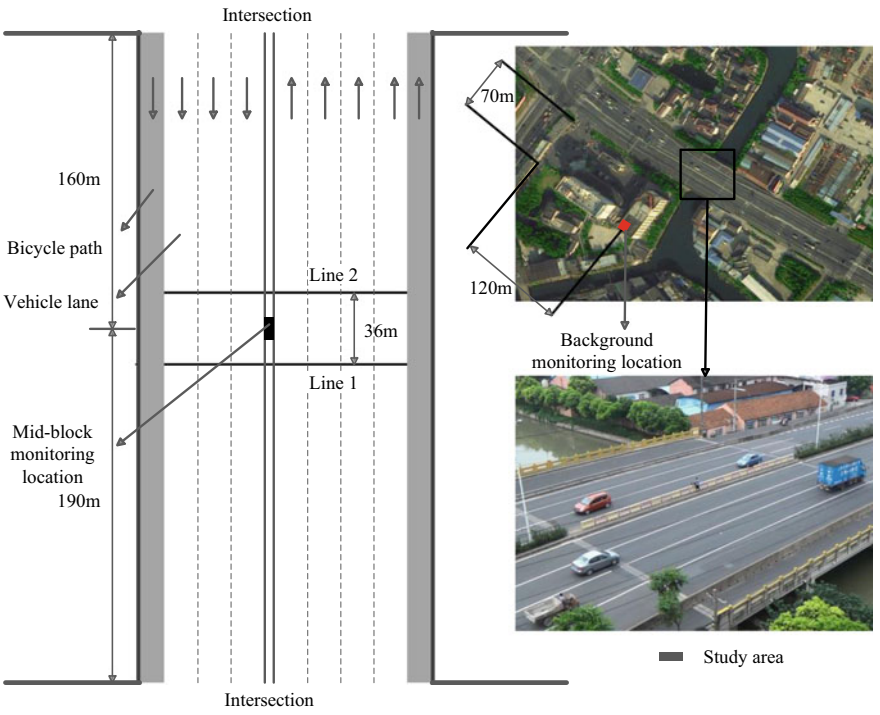


Fig. 1 Layout of the data investigation site

physically separated from bicycle lanes without interruption from pedestrians and non-motorized vehicles. The observation site is located in the mid-block with 160 m to the downstream intersection and 190 m to the upstream intersection.

To traffic conditions, traffic volume can be counted through a certain road section in ten seconds. The vehicles were separated into four types, named passenger car, bus, light-duty truck and heavy-duty truck. Bus and heavy-duty truck can be recorded as heavy-duty vehicle. The heavy-duty vehicle fraction can be calculated based on the different types of traffic volume. To the traveling speed, the moments that every vehicle crossed the two lines were recorded.

To $PM_{2.5}$ concentrations, two digital dust meters from TSI Incorporated were used to record the mid-block and background $PM_{2.5}$ concentrations second by second. The type of these two meters is DUSTTRAK TM DRX 8533. Video camera was placed on a roof with seven floors. One important thing is that time to collect traffic information and the $PM_{2.5}$ concentrations should be consistent. Data was recorded during a weekday from 16:30 to 18:30, with 15 min worming time period, which covers peak and non-peak traffic periods. Video tapes were later reviewed to obtain the traffic conditions.

2.2 Data Preparation

2.2.1 Analysis of Mid-block and Background $PM_{2.5}$ Concentrations

Data of $PM_{2.5}$ concentrations were collected by two digital dust meters second by second and also handled with continuous ten seconds intervals, as shown in Fig. 2. Background $PM_{2.5}$ concentration was less than the mid-block $PM_{2.5}$ concentration. The average values of mid-block and background $PM_{2.5}$ concentrations were 222.46

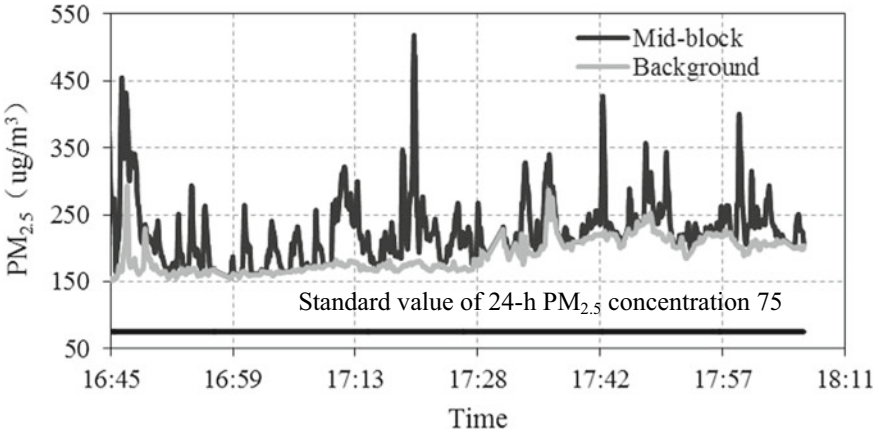


Fig. 2 Data of average $PM_{2.5}$ concentrations for continuous ten seconds

and 190.92 ug/m^3, and the two data sets have significant difference based on the result of t-test with the p-value (0.00) less than 0.05. Besides, the values of background $PM_{2.5}$ concentration were the least values of mid-block $PM_{2.5}$ concentration at a certain moment. The reason is that values of $PM_{2.5}$ concentrations showed consistent without the influence of traffic vehicles. It means that the difference between mid-block and background $PM_{2.5}$ concentrations resulted in the different traffic conditions. Furthermore, the standard value of 24-h $PM_{2.5}$ concentration in China is 75 ug/m^3, half of the background $PM_{2.5}$ concentration. Measures must be taken to decrease the $PM_{2.5}$ concentration.

2.2.2 Data of Micro Traffic Conditions

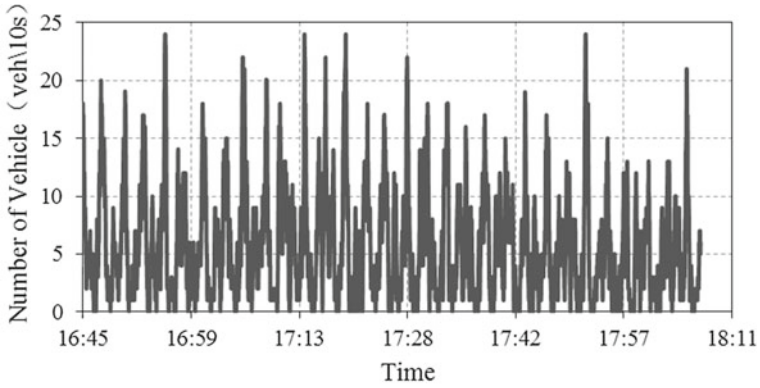
A video software developed by our research group was used to exact data from the video tapes. We can record the moment of every vehicle when passing line 1 and line 2 as shown in Fig. 1. Our study adopted continuous ten second to calculate the traffic volume and heavy-duty vehicle fraction, as shown in Fig. 3. The base line was line 1.

To traffic volume, the total number of bidirectional vehicles during continuous ten seconds presented a phenomenon of fluctuation. The maximum of traffic volume was 24 veh/10 s, while the minimum was 0 veh/10 s because of the signal control of the upstream and downstream intersections, and the average traffic volume is about 6 veh/10 s. The proportion of heavy-duty vehicle also fluctuated violently, from 0 to 100%; the average value of heady-duty vehicle fraction is about 18%.

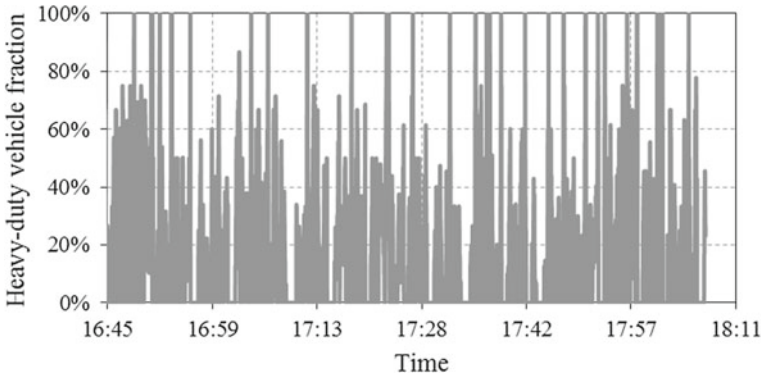
Besides, traveling speed of every vehicle can be calculated as the following formula:

$$v = \frac{36}{|t_2 - t_1|} \quad \text{(m/s)} \tag{1}$$

where t_2 means the time of vehicle passing line 2; t_1 means the time of vehicle passing line 1; the length between line 2 and line 1 is 36 m. Then traveling speed of every different types of vehicle could be recorded, as shown in Table 1. The average speed of passenger car was the largest, with the value of 46.49 km/h; while the others were 42.01 km/h, 37.84 km/h and 34.10 km/h respectively. Because of traffic congestion, the minimum speeds of the four types of vehicle were 7.04 km/h, 17.75 km/h, 7.24 km/h and 12.06 km/h, respectively. To the maximum speed, the range of different types of vehicle showed from 46.69 km/h (heavy-duty truck) to 79.63 km/h (passenger car). Overall, data of traffic information presented variety of different micro traffic conditions, which satisfied the data requirement of this study's objective.



(a) Traffic volume



(b) Heavy-duty vehicle fraction

Fig. 3 Data of traffic volume and heavy-duty vehicle fraction

Table 1 Traveling speed of different types of vehicle

Index	Passenger car	Bus	Light-duty truck	Heavy-duty truck
Average speed (km/h)	46.49	42.01	37.84	34.10
Minimum speed (km/h)	7.04	17.75	7.24	12.06
Maximum speed (km/h)	79.63	77.26	63.94	46.69
Standard deviation (km/h)	8.44	11.06	9.59	6.48

3 Impact of Traffic Congestion on Mid-block $PM_{2.5}$ Concentrations

3.1 Impact of Traffic Volume

Data of $PM_{2.5}$ concentrations and traffic information presented features of time series, fluctuating with time. Method of time series was adopted to illustrate the significant influence of traffic volume on $PM_{2.5}$ concentrations.

Autocorrelation function (ACF) and a partial autocorrelation function (PACF) plots of traffic volumes can be used to illustrate the cyclic nature of vehicle presence because of the upstream and downstream intersections. ACF and PACF plots show the similarities between observations as a function of the time lags between the observations. Given lag h, the ACF does not account for linear dependence between time t and time $t + h$; the PACF removes the linear dependence for the observations at time $t + 1$ through time $t + h - 1$, indicating the unique autocorrelation for lag h. Both of these functions are commonly used as a method for determining repeating patterns (7).

Cyclical arrival times were facilitated by upstream signals in either direction. The south upstream intersection had a cycle length of 160 s during the data collection period, and the north intersection had a cycle length of 90 s. These cycles are evident in the vehicle volume ACFs and PACFs in Fig. 4, where one lag equals 10 s. Directional

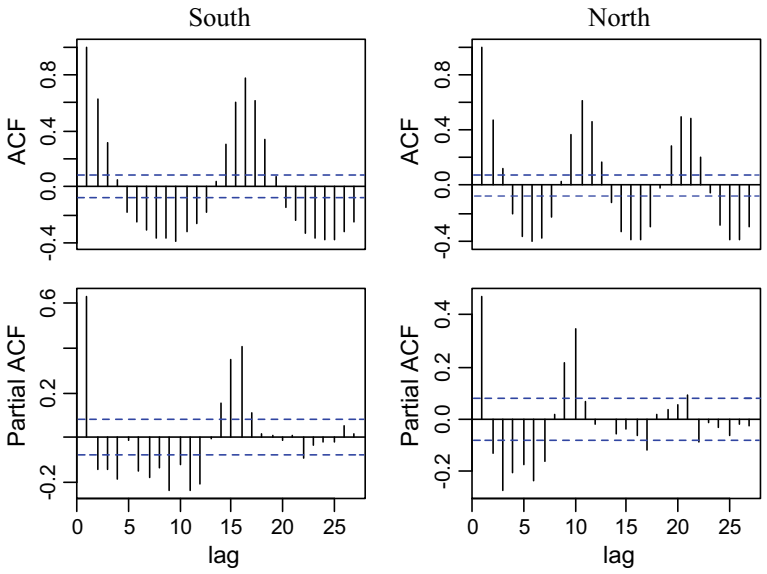


Fig. 4 Autocorrelation function of vehicle volumes showing platooning with upstream and downstream signal cycles (1 lag = 10 s)

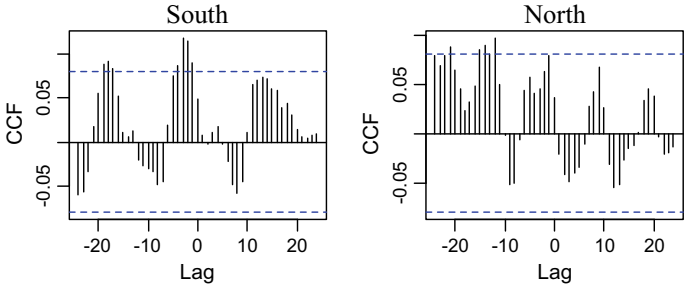


Fig. 5 Cross-correlation of vehicle volumes with $PM_{2.5}$ concentrations (1 lag = 10 s)

differences in the ACFs and PACFs are evident. The south direction response has a clear spike at 160 s; while to the north direction, the clear spike locates at 100 s.

The cyclical arrival times are referred to as the grouping of vehicles after departing from upstream and downstream intersections. To investigate the effect of vehicle traffic volume on $PM_{2.5}$ concentrations, a cross-correlation function (CCF) was made for each direction, as shown in Fig. 5.

The CCF in Fig. 5 presented the sample correlations between $PM_{2.5}$ concentrations at time t and traffic volumes at time $t + h$ for $h = 0, \pm1, \pm2, \pm3$, etc. Negative values for h indicated a correlation between volumes at a time h units before t and $PM_{2.5}$ concentrations at time t. The dashed lines in Fig. 5 showed the statistical significance level, calculated using:

$$\frac{Z_1 - \frac{\alpha}{2}}{\sqrt{N}} \tag{2}$$

where Z is the Z-value at a given level of significance and N means the sample size. At 95% confidence, the significance threshold is 0.078. To the south direction of traffic flow, $PM_{2.5}$ concentrations were significantly positively correlated (+9.1%) with vehicles passing at 18 lags (180 s). And to the north direction of traffic flow $PM_{2.5}$ concentrations were also significantly positively correlated (+9.7%) with vehicles passing at 12 lags (120 s). This lag times roughly matched the upstream and downstream cycle lengths, and the ACFs and PACFs. Therefore, traffic flow was significantly positively correlated with $PM_{2.5}$ concentrations.

3.2 Traffic Congestion and Mid-block $PM_{2.5}$ Concentrations

Based on the Code for Design of Urban Road Engineering in China [17], the conventional passenger car equivalents to the four types of vehicle are 1, 2, 1.5 and 2.5. Saturation was calculated with ten seconds intervals based on the convert coefficient. Table 2 illustrates the classification standard of different Level of Service (LOS).

Table 2 The classification standard of different LOS

LOS	A	B	C	D	E	F
Volume/capacity	≤0.27	0.27–0.57	0.57–0.70	0.70–0.85	0.85–1.00	>1.00

Figure 6 shows boxplots of mid-Data of PM2.5 concentrations segmented by traffic LOS. The boxplots show the range, upper/lower quartiles, and median observed values, with statistical outliers as red crosses. Figure 6 also includes the number of continuous 10-s aggregation intervals included in the plot for each LOS (as "N").

As can be seen in Fig. 6, mid-block $PM_{2.5}$ concentrations do not notably trend up or down with different LOS. Using a non-parametric Wilcoxon signed-rank test to compare each LOS in Fig. 6 with its neighbors, the LOS D versus LOS E and LOS E versus LOS F comparisons are statistically significantly different at $p = 0.02$ and $p = 0.004$ respectively. To LOS D versus LOS E, Observe that the difference in medians is small compared to the range of concentrations observed. To LOS E versus LOS F, the difference is lower concentrations at the heavier congestion level. Therefore, mid-block $PM_{2.5}$ concentrations are not correlated with traffic LOS. The

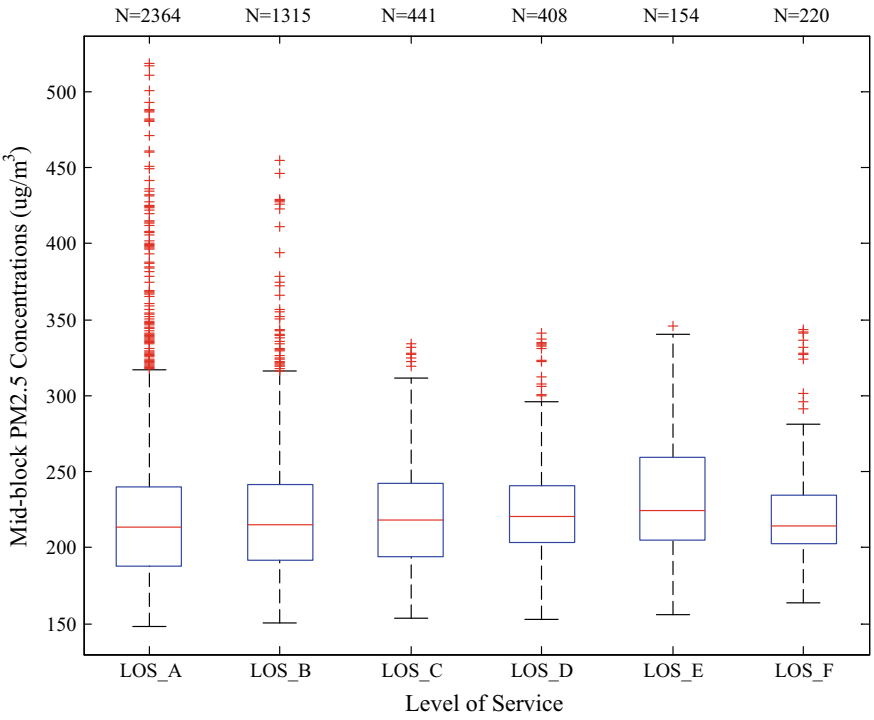


Fig. 6 Comparisons of traffic LOS and mid-block $PM_{2.5}$ concentrations (10 s aggregated data)

relationship between traffic conditions and $PM_{2.5}$ concentrations is explored in more detail using multiple regressions analysis below in Sect. 4.

4 Modeling of Micro Traffic Conditions on $PM_{2.5}$ Concentrations

4.1 Interpretation of Variables

Variables considered for the modeling of traffic conditions are listed in Table 3. To mitigate the influence of meteorological elements, the difference between mid-block road $PM_{2.5}$ concentrations and background $PM_{2.5}$ concentrations can be considered mainly caused by micro traffic conditions.

Furthermore, the function to calculate vehicle emission factor is an exponential function [18], as follows:

$$e = \exp\left(\sum_{i=0}^{4}\left(a_i \times v^i\right)\right) = \exp\left(a_0 + a_1 \times v + a_2 \times v^2 + a_3 \times v^3 + a_4 \times v^4\right) \quad (3)$$

where a_i are fitted parameters, v^i means the vehicle's speed of i power, e means emission factor (g/km). Therefore, the natural log of average concentration was used as dependent variable. It means that

Table 3 Variable definitions for modeling of the traffic conditions

Variable	Description	Unit
$ln(PM_{2.5})$	Natural log of average concentrations during ten seconds	ug/m^3
$M_{pm2.5}$	Mid-block road $PM_{2.5}$ concentrations during ten seconds	ug/m^3
$B_{pm2.5}$	Background $PM_{2.5}$ concentrations during ten seconds	ug/m^3
Vol_{pc}	Volume of passenger car during ten seconds	veh/10 s
$Speed_{pc}$	Average speed of passenger car during ten seconds	m/s
Vol_b	Volume of bus during ten seconds	veh/10 s
$Speed_b$	Average speed of bus during ten seconds	m/s
Vol_{ldt}	Volume of light-duty truck during ten seconds	veh/10 s
$Speed_{ldt}$	Average speed of light-duty truck during ten seconds	m/s
Vol_{hdt}	Volume of heavy-duty truck during ten seconds	veh/10 s
$Speed_{hdt}$	Average speed of heavy-duty truck during ten seconds	m/s
$Speed_{tf}$	Traffic flow speed during ten seconds	m/s
Vol_m	Modified total volume during ten seconds	mpcu/10 s
Occ_{hdv}	Heavy-duty vehicle fraction during ten seconds	%

$$ln(PM_{2.5}) = ln\left(M_{pm2.5} - B_{pm2.5}\right) \tag{4}$$

Traffic variables included volume of different types of vehicle and speed of different types of vehicle. The following formula was adopted to calculate the traffic flow speed:

$$Speed_{tf} = \frac{\sum Vol_i \times Speed_i}{\sum Vol_i} \quad (i = pc, b, ldt, hd)) \tag{5}$$

To total volume of traffic flow, previous studies were focused on road capacity; while in this study, passenger car equivalents were modified from the aspect of vehicles' different contributions to $PM_{2.5}$ concentrations. Besides, variable of heavy-duty vehicle fraction was also considered.

4.2 Modified Passenger Car Equivalents

From the aspect of vehicle emissions, researchers suggested that passenger car equivalents should reflect the different weight of different types of vehicle to different air pollutions. In this study, the multiple linear regressions model was established to modify the passenger car equivalents from the aspect of contribution on $PM_{2.5}$ concentration.

$$\ln(PM_{2,5}) = \sum b_i \times Vol_i + c \times Speed_{tf} + d \quad (i = pc, b, ldt, hdt) \tag{6}$$

where b_i, c, d are fitted parameters, $i = 1, 2, 3, 4, 5$.

Results were shown in Table 4. Overall, the adjust R-square was 0.5154. All the variables except Vol_{pc} have significant influence on the $PM_{2.5}$ concentration. To variable of Vol_{pc}, the reason is that passenger cars are mostly gasoline cars with little emission of $PM_{2.5}$, resulting in the lowest coefficient, with the value of 0.0157. The value of Vol_{hdt} was the largest, with the value of 1.32, and value of Vol_b following closely (0.82). And the coefficient of Vol_{ldt} was 0.287. Therefore, the modified passenger car equivalents were 1:52:18:84, with a big difference from the aspect of road

Table 4 Results of modified passenger car equivalents

Variable	Parameter estimate	Standard error	t value	Pr > |t|
d	5.88355	0.12279	47.92	<0.0001
Vol_{pc}	0.01568	0.02174	0.72	0.4715
Vol_b	0.82201	0.16992	4.84	<0.0001
Vol_{ldt}	0.28747	0.11472	2.51	0.0130
Vol_{hdt}	1.32014	0.30329	4.35	<0.0001
$Speed_{tf}$	−0.12559	0.00853	−14.72	<0.0001

Table 5 Results of the comprehensive model

Variable	Parameter estimate	Standard error	t value	Pr > |t|
g	5.76266	0.11307	50.97	<0.0001
Vol_m	0.01396	0.00207	6.75	<0.0001
Occ_{hdv}	−0.25986	0.06100	−4.26	<0.0001
$Speed_{tf}$	−0.09086	0.01127	−8.06	<0.0001

capacity. Some studies pointed out that the ration on emission factor of $PM_{2.5}$ to bus, light-duty truck and heavy-duty truck was 3:1:5, consistent with our modified passenger car equivalents [19]. Besides, the coefficient of $Speed_{tf}$ was negative; indicating that improving traveling speed and alleviating traffic congestion can decrease the $PM_{2.5}$ concentration effectively.

4.3 A Comprehensive Model

Based on the above analysis, a comprehensive model was established to explore the contribution of traffic conditions on $PM_{2.5}$ concentration. There exists correlation between variable of Vol_m and variable of Occ_{hdv}, so the comprehensive model is as follows:

$$\ln(PM_{2,5}) = f_1 \times Vol_m + f_2 \times \frac{1}{Occ_{hdv}} + f_3 \times Speed_{tf} + g \tag{7}$$

where f_i, g are fitted parameters, $i = 1, 2, 3$.

Table 5 presented the results of fitted parameters. Overall, the adjust R-square was 0.5567. Results showed that with the increase of total volume, the $PM_{2.5}$ concentration could increase. And the heavy-duty vehicle fraction presented a positive effect. When there are more and more heavy-duty vehicles, the $PM_{2.5}$ concentration will be increasing rapidly. Therefore, many cities have promulgated a decree that heavy-duty trucks are prohibited passing through cities' road during daytime. Furthermore, low traffic flow speed can result in high $PM_{2.5}$ emission factor, increasing $PM_{2.5}$ concentration. Besides, because of fugitive dust, the value of constant term was 5.76, with 20% of the average $PM_{2.5}$ concentration (31.54 ug/m^3).

5 Conclusions

This study aimed to evaluate the traffic congestion on mid-block $PM_{2.5}$ concentrations. Based on the field observations, mid-block and background $PM_{2.5}$ concentrations were recorded to calculate the $PM_{2.5}$ concentration influenced by different

traffic conditions. Traffic volume, heavy-duty vehicle fraction and traveling speed were extracted from video tapes at ten seconds intervals. Relationship between traffic congestion and mid-block $PM_{2.5}$ concentrations was analyzed. A modified passenger car equivalent was proposed from the aspect of contribution on $PM_{2.5}$ concentrations. At last, a comprehensive model was established to evaluate the impact of micro traffic conditions on $PM_{2.5}$ concentrations.

Results showed that traffic volume had significant influence on mid-block PM2.5 concentrations. Mid-block PM2.5 concentrations were not correlated with traffic LOS. The multiple linear regressions model was established to modify the passenger car equivalents, with the value of 1:52:18:84. Results of the comprehensive model showed with the increase of total volume or heavy-duty vehicle fraction, the $PM_{2.5}$ concentrations could increase. Besides, low traffic flow speed can result in high $PM_{2.5}$ emission factor, leading to the increase of $PM_{2.5}$ concentrations.

The findings of this study can help better understand the traffic congestion and micro traffic conditions on $PM_{2.5}$ concentrations. More importantly, findings of this study can be helpful to police makers. First, the passenger car equivalent of heavy-duty truck to $PM_{2.5}$ is about 84, that is why many cities have promulgated a decree that heavy-duty trucks are prohibited passing through cities' road during daytime. Second, the passenger car equivalent of bus to $PM_{2.5}$ is about 52, resulting in high $PM_{2.5}$ concentration around bus stations. Third, improving traveling speed and alleviating traffic congestion can decrease the $PM_{2.5}$ concentration effectively. Last but not least, the comprehensive model can be used to estimate $PM_{2.5}$ concentrations under the different traffic conditions.

Foundation Item Project supported by the Fundamental Research Funds for the Central Universities (No. 2018B08014) and by the National Science Foundation of China (No. 51608171).

References

1. Klemm RJ, Mason RM, Heilig CM, Neas LM, Dockery DW (2000) Is daily mortality associated specifically with fine particles? Data reconstruction and replication of analyses. J Air Waste Manage Assoc 50(7):1215–1222
2. Ambient Air Quality Standards (2012) Ministry of Environmental Protection, P.R. China
3. National Ambient Air Quality Standards (NAAQS) (2012) Environmental Protection Agency, United States of America
4. Three Factors Result in the Haze of Beijing, Local Pollutant Emission is Top the List. Chinese Radio Network. http://news.cnr.cn/native/gd/201411/t20141104_516722367.shtml. Accessed 4 Nov 2014
5. Results of Atmospheric Particulates Source in Shanghai. Chinese Environment Network. http://www.cenews.com.cn/xwzx2013/hjyw/201501/t20150109_786238.html. Accessed 9 Jan 2015
6. Vehicle Emission of PM2.5 is about 24.6% in Nanjing. Yangzi Evening News Network. http://www.21cs.cn/details/?id=528126. Accessed 30 Apr 2015
7. Moore A, Figliozzi M, Bigazzi A (2014) Modeling impact of traffic conditions on variability of midblock roadside fine particulate matter case study of an urban arterial corridor. Transp Res Rec 2428:35–43

8. Wang Z, He H, Lu F, Lu QC, Peng ZR (2015) A hybrid model for prediction of carbon monoxide and fine particulate matter concentrations near road intersection. Transp Res Rec 2503:29–38
9. Cai M, Yin Y, Xie M (2009) Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. Transp Res Part D 14:32–41
10. Wang Y, Li J, Cheng X, Lun X, Sun D, Wang X (2014) Estimation of PM10 in the traffic-related atmosphere for three road types in Beijing and Guangzhou, China. J Environ Sci 26:197–204
11. He H, Lu WZ, Xue Y (2009) Prediction of PM10 concentrations at urban traffic intersections using semi-empirical box modelling with instantaneous velocity and acceleration. Atmos Environ 43:6336–6342
12. Liu H, Chen X, Wang Y, Han S (2013) Vehicle emission and near-road air quality modeling for Shanghai, China: based on global positioning system data from taxis and revised moves emission inventory. Transp Res Rec 2340:38–48
13. Su F, Roorda MJ, Miller EJ, Morrow E (2015) An integrated approach to estimate pedestrian exposure to roadside vehicle pollutants. In: TRB 94th annual meeting, Washington D.C.
14. Chai M, Wei H, Li Z, Lu M (2010) Modeling impact of traffic operation on carbon monoxide dispersion. In: ICCTP: integrated transportation systems-green intelligent reliable, pp 2890–2902
15. Venkatram A, Isakov V, Seila R, Baldauf R (2009) Modeling the impacts of traffic emissions on air toxics concentrations near roadways. Atmos Environ 43:3191–3199
16. Pu Y, Yang C (2014) Estimating urban roadside emissions with an atmospheric dispersion model based on in-field measurements. Environ Pollut 192:300–307
17. Code for Design of Urban Road Engineering (2012) Ministry of Housing and Urban-Rural Development, P.R. China
18. Bigazzi AY, Figliozzi MA (2012) Congestion and emissions mitigation: a comparison of capacity, demand, and vehicle based strategies. Transp Res Part D 17:538–547
19. Huang J, Liu Y, Cheng X, Li H, Lai D (2014) Vehicle emission characteristics of PM2.5 with COPERT IV model. Environ Sci Technol 37(1):43–47

Uncertainty Analysis of Rock Strength Based on Mohr-Coulomb Criterion

Yongfeng Ma and Rangang Yu

Abstract Due to uncertainties of the input parameters, such as the maximum principle stress, the minimum principle stress, the cohesion and the internal friction angle, the evaluation of the rock strength becomes the uncertainty problem. In such case, the uncertainty analysis method was proposed to deal with the uncertainty of the rock strength based on Mohr-Coulomb criterion. To describe the properties of the random variables, the indoor experiment was performed to obtain the mean and the standard deviation of the input parameters. Furthermore, the random technique was introduced to yield the sample data based on the mean and the standard deviation. More importantly, the factor of safety (FOS) was defined to evaluate the rock strength based on Mohr-coulomb criterion. Meanwhile, Monte Carlo method was introduced to the probability of the factor of safety. The results show that FOS has a range of [0.6, 1.4], and shows an obvious normal distribution.

Keywords Rock strength · Uncertainty · Monte Carlo method · Mohr-Coulomb criterion

1 Introduction

Rock strength evaluation is a basic objective for the practice engineering. Till now, different criteria different were established to evaluate the rock strength, such as Mohr-Coulomb criterion [16], the modified Lade criterion [6, 13], Drucker-Prager criterion [5, 14], Hoek-Brown criterion [8], Mogi-Coulomb criterion [1, 15]. Using these criteria, the rock strength was evaluated deterministically. However, the uncertainty of the parameter leads to the uncertainty of the rock strength [2, 4, 7, 17, 20]. In view of this, the uncertainty of the rock strength should be considered seriously.

Y. Ma (✉) · R. Yu
College of Pipeline and Civil Engineering, China University of Petroleum, Qingdao 266580, China
e-mail: yongfeng314@126.com

Y. Ma
China Petroleum Engineering & Construction Corp., Beijing 100120, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_107

Currently, some researchers have developed the corresponding models to address the uncertainty of the rock strength. Bozorgzadeh et al. [3] established a reliability-based design in rock engineering. Jia et al. [10] proposed reliability theory should be used to analyze the shearing stability of CSGR dams, considering the uncertainty of the shearing parameters. Langford and Diederichs [12] quantified the uncertainty in Hoek–Brown intact strength envelopes. Hoek [9] evaluated the rock mass properties based on reliability. Yang et al. [19] used the response surface method (RSM) and Monte-Carlo simulation to calculate the Hasofer-Lind reliability index and the failure probability. Wu et al. [18] used a reliability-based approach to evaluate the stability of high rockfill dams based on a nonlinear shear strength criterion. Langford and Diederichs [11] proposed the reliability based approach to tunnel lining design using a modified point estimate method.

In this paper, the uncertainty analysis method was proposed to examine the uncertainty of the rock strength based on Mohr-Coulomb criterion, considering the uncertainty of the random variable. Especially, the factor of safety was defined to evaluate the rock strength. Finally, Monte Carlo method was introduced to calculate the probability on the factor of safety.

2 Uncertainty Analysis of Input Parameters

2.1 Experiment Evaluation of Uncertainty

Through the indoor experiment, the rock mechanical parameters were obtained and shown in Table 1, including the maximum principle stress, the minimum principle stress, the cohesion and the internal friction angle.

According to the experiment data, the properties of the input parameter were calculated and listed in Table 2.

Table 1 Rock mechanical parameters

Rock samples	σ_1 (MPa)	σ_3 (MPa)	c (MPa)	φ (°)
R-1	88.53	21.32	8.48	32.0
R-2	90.12	17.88	7.19	30.6
R-3	84.22	18.03	7.02	26.8
R-4	82.64	19.72	8.02	28.7
R-5	80.78	20.11	7.55	27.0
R-6	78.02	20.53	7.12	30.5
R-7	76.39	18.36	6.93	31.2
R-8	75.44	17.75	7.44	28.3
R-9	86.36	16.99	6.68	26.4
R-10	85.26	20.10	7.52	29.5

Table 2 Properties of the input parameter

Random variables	Maximum value	Minimum value	Mean	Uncertainty in estimation (%)	Standard deviation
σ_1 (MPa)	90.12	75.44	82.78	+8.87/−8.91	5.05
σ_3 (MPa)	21.32	16.99	19.08	+11.74/−10.95	2.09
c (MPa)	8.48	6.68	7.40	+14.59/−9.73	0.54
φ (°)	32	26.4	29.1	+9.97/−9.28	1.97

2.2 *Probability Distribution of Input Parameters*

After obtaining the mean and the standard deviation of parameters, the random sampling technique was used to create plenty of samples. In this analysis, Monte Carlo method, a classic sampling technique, was introduced to generate 100,000 sample data. These samples were used to fit the probability curves of the random variables, seeing Figs. 1, 2, and 3.

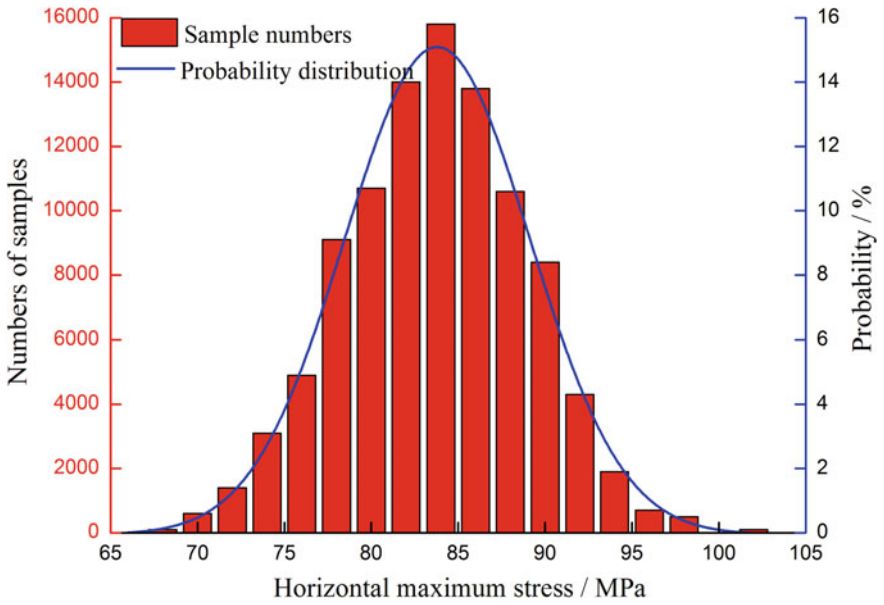


Fig. 1 Uncertainty quantification for the horizontal maximum stress

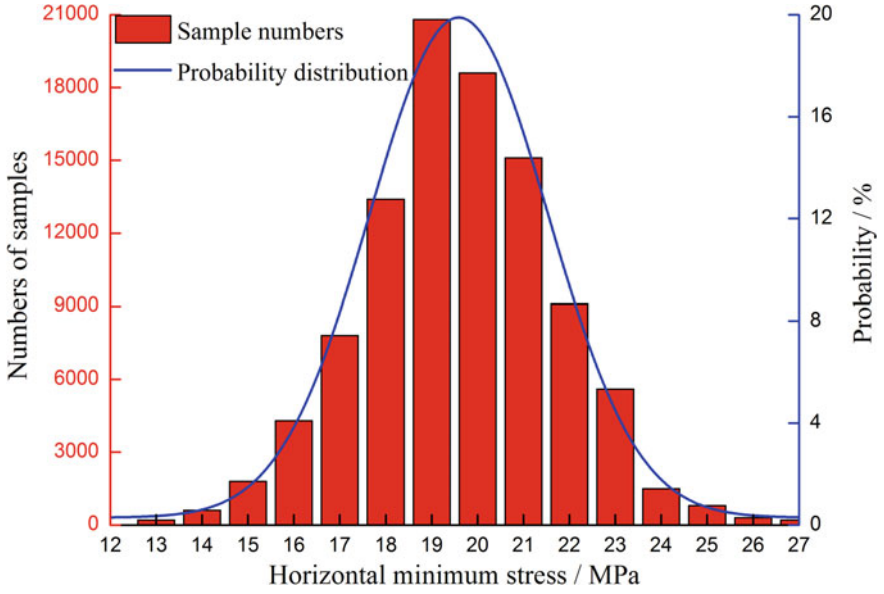


Fig. 2 Uncertainty quantification for the horizontal minimum stress

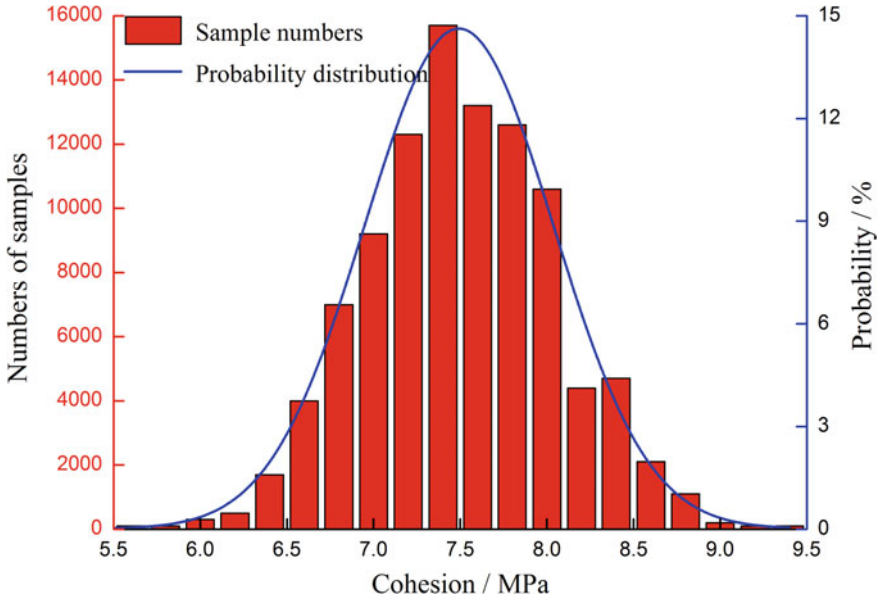


Fig. 3 Uncertainty quantification for the cohesion

3 Uncertainty Analysis of Rock Strength

3.1 Mohr-Coulomb Criterion

Currently, Mohr-Coulomb criterion was widely used to evaluate the rock strength, seeing Eq. (1):

$$\sigma_1 - A\sigma_3 = B \tag{1}$$

where $A = \frac{1+\sin\varphi}{1-\sin\varphi}$, $B = \frac{2c\cos\varphi}{1-\sin\varphi}$, σ_1 is the maximum principle stress, σ_3 is the maximum principle stress, φ is the internal friction angle, c is the cohesion.

3.2 Evaluation Index

In this section, the factor of safety (FOS) was selected as an evaluation index to determine the rock strength. FOS was defined as Eq. (2):

$$\mathrm{FOS} = \frac{\sigma_3 \frac{1+\sin\varphi}{1-\sin\varphi} + \frac{2c\cos\varphi}{1-\sin\varphi}}{\sigma_1} \tag{2}$$

FOS > 1 means that rock strength satisfies the engineering requirement. Otherwise, the rock strength does not satisfy the engineering requirement.

3.3 Probability Prediction

Monte Carlo method is based on the random number, to yield the sample data. Assuming that the total of n sample data were generated, the mean and the standard deviation can be calculated based on the following Eqs. (3) and (4).

$$\mu_{\mathrm{FOS}} = \sum_{i=1}^{n} (\mathrm{FOS})_i / n \tag{3}$$

$$\sigma_{\mathrm{FOS}} = \sqrt{\frac{\sum_{i=1}^{n} [(\mathrm{FOS})_i - \mu_{\mathrm{FOS}}]^2}{n-1}} \tag{4}$$

Theoretically, the mean and the standard deviation can converge to the accurate solution as long as the sample data were adequate, when using Monte Carlo method.

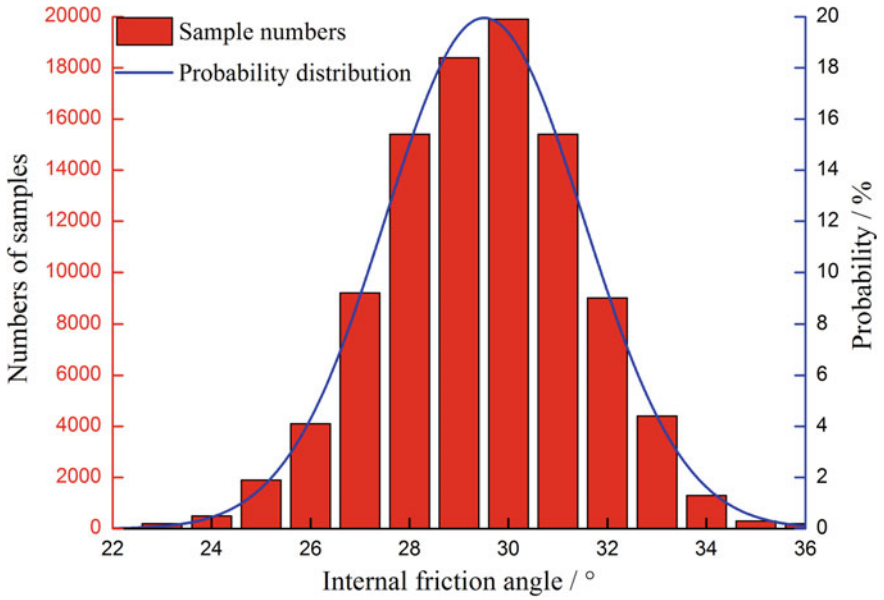


Fig. 4 Uncertainty quantification for the internal friction angle

4 Results and Discussions

4.1 Result Discussion

Combing Figs. 1, 2, 3 and 4, the uncertainty of rock strength can be evaluated, considering the uncertainties of the maximum principle stress, the minimum principle stress, the cohesion and the internal friction angle. The results were shown in Fig. 5.

According to the result, FOS has a range of [0.6, 1.4], which means that rock strength is in the unstable state. Additionally, the probability on FOS can be observed clearly. At FOS = 0.98, the probability is maximum, equaling to 3.5%. More importantly, FOS shows an obvious normal distribution.

5 Conclusions

(1) In essence, evaluation of the rock strength belongs to the uncertainty problem, which results from the uncertainty of the input parameter. In view of this, uncertainty of the input parameters was considered to evaluate the rock strength more accurately, including the maximum principle stress, the minimum principle stress, the cohesion and the internal friction angle.

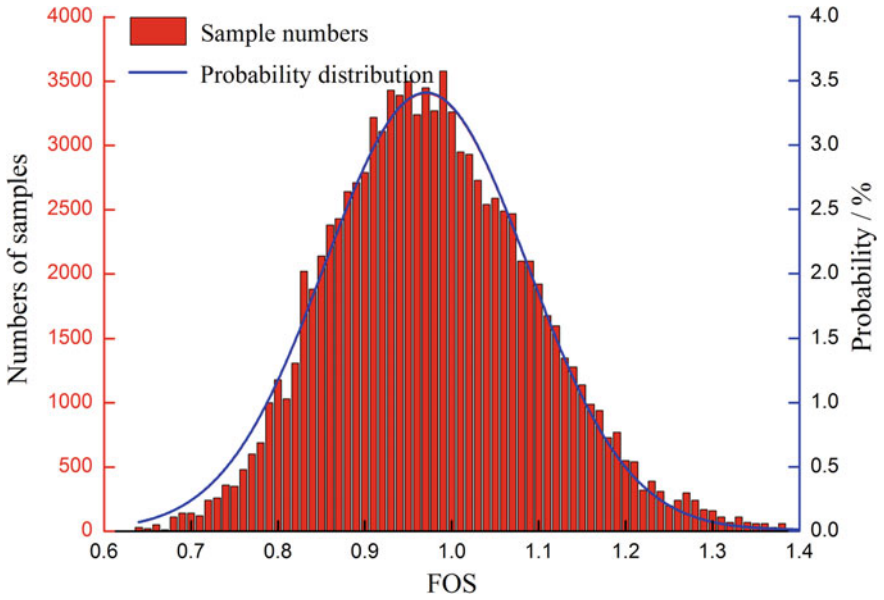


Fig. 5 Uncertainty quantification for rock strength based on Monte Carlo method

(2) Mean and standard deviation of input parameters play the important role for the uncertainty quantification of the rock strength. In view of this, the indoor experiment was carried out to obtain the mean and the standard deviation of the input parameters. Furthermore, Monte Carlo method was used to generate 100,000 sample data based on the mean and the standard deviation.

(3) To evaluate the rock strength, the factor of safety (FOS) was defined based on Mohr-coulomb criterion. Meanwhile, Monte Carlo method was introduced to the probability of the factor of safety. The results show that FOS has a range of [0.6, 1.4]. At FOS = 0.98, the probability is maximum, equaling to 3.5%. More importantly, FOS shows an obvious normal distribution.

Acknowledgements The author is very much indebted to the Projects Supported by Fundamental Research Funds of China Petroleum Engineering & Construction Corp (No. CPGEC2017KJ01) for the financial support.

References

1. Al-Ajmi AM, Zimmerman RW (2006) Stability analysis of vertical boreholes using the Mogi-Coulomb failure criterion. Int J Rock Mech Min Sci 43:1200–1211
2. Al-Ajmi AM, Al-Harthy MH (2010) Probabilistic wellbore collapse analysis. J Pet Sci Eng 74:171–177

3. Bozorgzadeh N, Escobar MD, Harrison JP (2018) Comprehensive statistical analysis of intact rock strength for reliability based design. Int J Rock Mech Min Sci 106:374–387
4. Colmenares LB, Zoback MD (2002) A statistical evaluation of intact rock failure criteria constrained by polyaxial test data for five different rocks. Int J Rock Mech Min Sci 39(6):695–729
5. Drucker DC, Prager W (1952) Soilmechanics and plastic analysis or limit design. Q Appl Math 10:157–165
6. Ewy RT (1999) Wellbore-stability predictions by use of a modified Lade criterion. SPE Drill Complet 14:85–91
7. Gholami R, Rabiei M, Rasouli V, Aadnoy B, Fakhari N (2015) Application of quantitative risk assessment in wellbore stability analysis. J Pet Sci Eng 135:185–200
8. Hoek E, Brown ET (1997) Practical estimates or rock mass strength. Int J Rock Mech Min Sci 34:1165–1186
9. Hoek E (1998) Reliability of Hoek-Brown estimates of rock mass properties and their impact on design. Int J Rock Mech Min Sci 35:63–68
10. Jia JS, Wang S, Zheng CY, Chen ZP, Wang Y (2018) FOSM-based shear reliability analysis of CSGR dams using strength theory. Comput Geotech 97:52–61
11. Langford JC, Diederichs MS (2013) Reliability based approach to tunnel lining design using a modified point estimate method. Int J Rock Mech Min Sci 60:263–276
12. Langford JC, Diederichs MS (2015) Quantifying uncertainty in Hoek-Brown intact strength envelopes. Int J Rock Mech Min Sci 74:91–102
13. Lade P, Duncan J (1975) Elastoplastic stress–strain theory for cohesionless soil. J Geotech Eng ASCE 101(10):1037–1053
14. Liu M, Gao Y, Liu H (2012) A nonlinear Drucker-Prager and Matsuoka-Nakai unified failure criterion for geomaterials with separated stress invariants. Int J Rock Mech Min Sci 50:1–10
15. Mogi K (1971) Fracture and flow of rocks under high triaxial compression. J Geophys Res 76(5):1255–1269
16. Mohr O (1900) Welcheumstande bedingendieelastizita tsgrenze unddenbruch eines materials? VDI-Z. 44:1524
17. Udegbunam JE, Aadnøy BS, Fjelde KK (2014) Uncertainty evaluation of wellbore stability model predictions. J Pet Sci Eng 124:254–263
18. Wu ZY, Li YL, Chen JK, Zhang H, Pei L (2013) A reliability-based approach to evaluating the stability of high rockfill dams using a nonlinear shear strength criterion. Comput Geotech 51:42–49
19. Yang XL, Zhou T, Li WT (2018) Reliability analysis of tunnel roof in layered Hoek-Brown rock masses. Comput Geotech 104:302–309
20. Yi X, Ong S, Russell JE (2006) Quantifying the effect of rock strength criteria on minimum drilling mud weight prediction using polyaxial rock strength test data. Int J Geomech 6(4):260–268

Dynamic Maintenance Decision Model for Essential Equipment of Metro Based on Markov Chain

Jun Wu and Yongneng Xu

Abstract Maintenance of essential equipment is an important factor in ensuring the safety operation of urban rail transit. The traditional maintenance decision-making usually has irrational repairs schedule and excessive maintenance. In order to proffer solution to the problems associated with the traditional maintenance decision-making, the current maintenance mode of equipment were analysed. After critical analysis of the traditional maintenance decision-making, based on Markov chain a Dynamic maintenance decision model and scientific planning of the decision-making process was proposed. Furthermore, the components of the essential equipment of rail transit was also analysed. Through preventive maintenance interval and frequency, reliability and other factors, a Markov chain prediction model is established to predict the time interval for repairing the equipment. The validity of the model was proved by processing the actual maintenance data with MATLAB, which provides a reasonable basis for the dynamic maintenance decision-making of equipment.

Keywords Markov chain · Essential equipment of the metro · Dynamic maintenance decision · Fault prediction

1 Introduction

The failure rate of most equipment is a function of time. The typical failure curve is called the Bathtub curve. The shape of the curve is high at both ends and low in the middle. It has obvious stages. It can be divided into three stages: early failure stage, accidental failure stage and serious failure stage. The traditional fault maintenance is based on the theory of "bathtub curve". Preventive maintenance and repairing maintenance are combined to ensure the maintenance of components. The traditional maintenance methods are easy to cause "over-maintenance" or "under-maintenance"

J. Wu · Y. Xu (✉)
Automated Institute, Nanjing University of Science and Technology, Nanjing, China
e-mail: x780906yn@163.com

J. Wu
e-mail: 1075542294@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_108

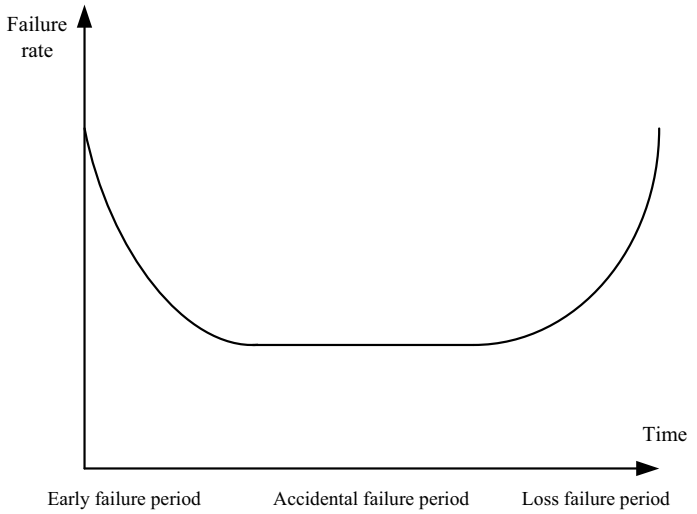


Fig. 1 Bathtub curve

[1]. Earliest state-based maintenance began in the 1970s. The so-called state-based maintenance is to obtain the state information of the system running from the sensors embedded in the equipment or the external testing equipment. Through real-time or near real-time evaluation of the state information, the maintenance requirements of the equipment are finally made, which makes the maintenance work more timely and accurate. At the same time, Save maintenance costs as much as possible. The bathtub curve is showed in Fig. 1.

With the continuous development and improvement of state-based maintenance methods, various countries have formed relatively complete standards for fault diagnosis, prediction and maintenance decisions. Cassady [2] in the United States has established three alternative maintenance models for complex structural systems, including minimum maintenance, repair of defective parts and preventive replacement parts. On this basis, imperfect maintenance activities are considered. Conroy et al. [3] designed an expert system that can perform fault diagnosis to achieve state-based maintenance decisions. Tomita [4] proposed a maintenance decision model based on vibration state changes. At present, in the maintenance guarantee of key equipment for rail transit, the state-based maintenance decision theory has been applied more widely, which can reduce maintenance costs and improve personnel efficiency.

Domestic research on the maintenance of subway equipment status is relatively late. In the research of equipment maintenance guarantee scheme, it only involves the qualitative description of the development and optimization of maintenance support scheme and the local simulation evaluation of the problem. It is rare to see the system maintenance guarantee scheme and optimize methods for quantitative research and related applications. Zhang [5] proposed a degraded state space segmentation method

for a two-component series system with clearly distinguishable fault characteristics. Based on the embedded Markov chain construction system semi-update period, the system long-term average cost rate model and the optimization solution system was established. The opportunity to repair the combined probability. Dong [6] conducted an in-depth study on the operation and maintenance decision support system of power generation equipment in his doctoral thesis. He proposed an improved analysis method to determine the operation and maintenance decision-making process of power generation equipment. The key technologies of equipment importance analysis, state evaluation and prediction, maintenance decision-making and optimization in the decision-making process are studied. Xin et al. [7] built a state maintenance decision model based on non-ideal maintenance, and discussed the improvement of equipment maintenance times. How to improve the reliability, optimize the maintenance methods used for maintenance, and improve the availability of equipment. The above research has made certain decisions on different equipments in various fields, and has achieved certain results. However, further research and discussion are needed on the dynamic mastery of the state, the relationship between the state transition and the actual state, and the guidance of the actual maintenance work.

Based on the above research, this paper uses the dynamic model based on Markov chain which selects the key equipment of the subway, adopting its annual historical maintenance data to get the actual state of the equipment. At the same time it obtains the status and failover probability for next moment. Which in turn to predict specific repair times, optimize maintenance plans and provide guidance for decision making.

2 Markov Chain Model Analysis

A. *Traditional maintenance mode*

The traditional rail transit equipment maintenance guarantee is mainly based on plan repair, that is, through regular inspection to find the problem and deal with it at first. And the main basic theory is P-F interval. Through the timing detection to determine the moment of the potential failure of the equipment, and take pre-maintenance or after-the-fact maintenance to protect. So as to avoid adverse effects of equipment failure [8, 9]. The earlier the potential fault is discovered, the longer the P-F interval. Long P-F intervals mean no frequent inspections or more time to take steps to avoid the consequences of failure. Therefore, maintenance staff spend a lot of effort to study the state of potential failures, and strive to develop techniques for detecting potential failures that provide the longest possible P-F interval.

Figure 2 shows that a long P-F interval means a potential fault is detected at a higher position on the P-F curve. However, the higher the curve rises, the smaller the deviation from the normal state. And the monitoring technique used to detect potential faults is more sensitive. Currently used is a condition monitoring technology that monitors the main parameters. Through the state monitoring technology, the analysis of certain parameters of the equipment such as vibration, temperature and

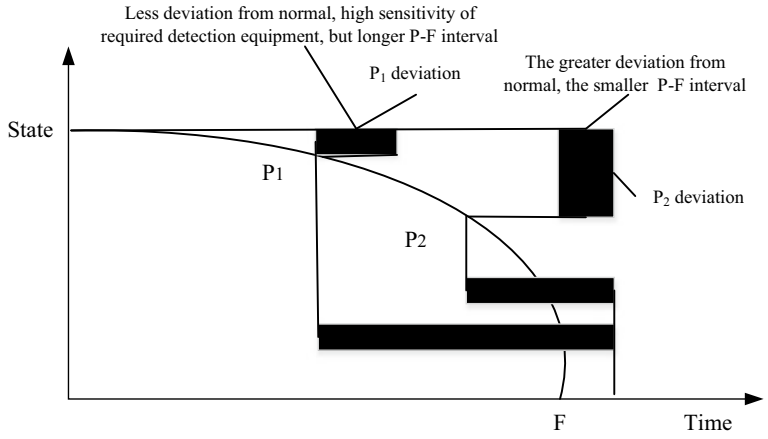


Fig. 2 Deviation between P-F interval and normal state

etc., to obtain the fault state of the equipment. So as to more clearly understand the fault law and accurately determine the P-F interval. Among them, in order to ensure the timely and effective maintenance of equipment, the P-F interval of the net surplus should be greater than the downtime of the equipment (That is, the maximum time required from finding potential faults to stopping the operation of the equipment.

This traditional maintenance method can easily cause "over-maintenance" or "under-maintenance" problems. which not only increases the production cost of the subway company, but also reduces the efficiency of employees and affects the normal operation of the train. In addition, when a fault occurs, it is impossible to locate the fault to a specific unit, which slows down the speed of maintenance support and is likely to cause serious accidents.

B. *Markov's Dynamic Maintenance Decision*

The process of based on Markov's decision-making mode of condition-based maintenance is that: firstly, the state of the object under test is detected and the characteristic quantity of the state information is obtained. Combining the historical status information and maintenance information of the equipment, the current status of the equipment is analyzed. The state interval and level of the equipment are judged at this time. Then, the time and probability of equipment failure are calculated by Markov chain transformation model and transfer matrix, and the preventive maintenance time is determined. The maintenance prediction results under this prediction mode provide technical support for higher prevention of potential failure.

By introducing the decision-making mode of condition-based maintenance based on Markov chain and changing the fixed maintenance interval of traditional planned maintenance, a new maintenance interval is proposed. Which not only improves the production efficiency of maintenance, but also ensures the normal operation of equipment.

Select the n technical status of the device, which are respectively recorded as S_1, S_2, …, S_n (where S_1 is the initial technical state of the device, that is the technical state when the device is enabled. And S_n is the technology state corresponding to the device when the device reaches the reliable life without failure). State K indicates that the device is in a fault state, and can also be regarded as the $n + 1$ technical state of the device.

Since the state of the device is always a transition from good to bad, it is a gradual and irreversible change process (except for man-made maintenance interventions), so the technical state transition of the equipment is a non-recurrent state transition. Assuming $i < j$, the probability of specifying state transition $S_i \rightarrow S_j$ is A_{ij}. If the device remains in the same fixed state for a continuous period of time, the probability of $S_i \rightarrow S_i$ is A_{ii}. The probability that non-faulty existence state of the device S_i to the fault state K is $A_i(1 - n) = (1 - P_i)$, where P_i is the reliability of the device in the S_i state [10, 11]. Then the constructed Markov chain is shown in Fig. 3.

Markov chain is introduced to predict the failure of key equipment: If the initial state of the equipment is S_i, the corresponding time is t_i. If the equipment is going to enter the failure state at this time, the technical state of the equipment that may experience includes S_i, S_{i+1},…, S_n. And all the technical states will eventually enter the absorbed state K, that is the failure state. Therefore, the time for predicting the transition of the device from normal state S_i to failure state K can be transformed into: the transition time of the device from state S_i to state S_{i+1}, …, S_n. That is the sum of time between all transition states.

Assuming that the state transition matrix of the device is A, the element on the position (i, j) is represented by $A(i, j) = A_{ij}$. By transposing the state matrix A of the equipment, it can be obtained

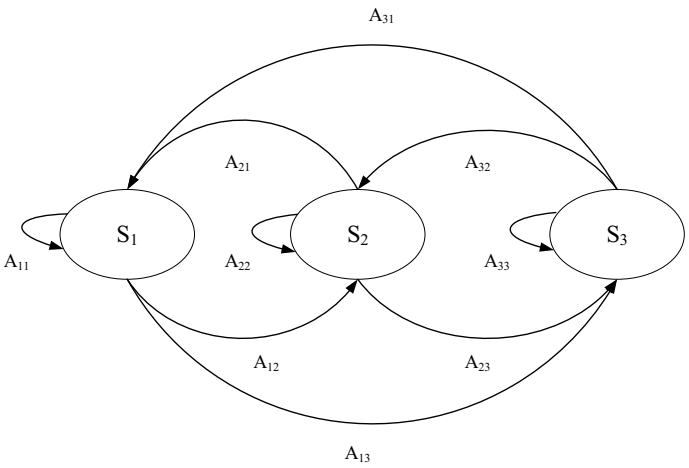


Fig. 3 Markov chain diagram

$$\overline{A} = \begin{bmatrix} \widetilde{A} & 0 \\ S & R \end{bmatrix}, \widetilde{A}^n = \begin{bmatrix} \widetilde{A}^n & 0 \\ S_n & R^n \end{bmatrix} \tag{1}$$

where: R means that the state transition matrix A of the device is transposed after removing the fault state K; the states in R are all non-fault states [12]. Let Y_j denote the number of times which device passes the non-faulty state S_j during the state transition, then there is

$$Y_j = \sum_{n=0}^{\infty} I\{X_n = S_j\}, \mathrm{I} = \begin{cases} 0, & Xn \neq S_j \\ 1, & X_n = S_j \end{cases} \tag{2}$$

Then the matrix I + R+R^2 + R^3 ⋯ the element at the (i, j) position is

$$\mathrm{E}\big(Y_j | X_0 = S_i\big) = E\left\langle \sum_{n=0}^{\infty} I\{X_n = S_j\} | X_0 = S_i \right\rangle = \sum_{n=0}^{\infty} A_n(i, j) \tag{3}$$

In this formula, A_n(i, j) represents the probability of being in a given state j after a state transition of n times from a certain state i. Let M = (I − R) − 1, from the nature of the matrix, $M = (I - R) - 1 = I + R + R^2 + R^3 + \cdots$. Then the element M_j at the position of (i, j) represents the average number of times that the state S_j is accessed from the state S_i.

We assume that: T_{n+1} is the average minimum time required for the device to reach the fault state K from the initial state; $T_{n+1,\sigma}$, (n + 1 ≠ σ) indicates that the device start time is in a fixed state, from this start state transition to the state K before the failure, the residence time in the state S_σ (non-fault state) (counted once when the initial state is S_σ). Then there is

$$\bar{T}_{n+1} = E(T_{n+1} | X_0 = S_i) = E\left\langle \sum_{\sigma \neq n+1} T_{n+1,\sigma} | X_0 = S_i \right\rangle = \sum_{\sigma \neq n+1} \Delta t_\sigma M_{i\sigma} \tag{4}$$

That is under the initial state S_i, the equipment to reach the fault state K through the state transition is the shortest average number of times which is required for the equipment. If the residence time of the equipment in each technical state can be known in advance, then the time required for the equipment to reach the fault state K can be calculated.

3 Maintenance Decision of Metro Key Equipment Based on Markov Chainsystem Classification

This paper selects the key equipment components of the metro: wheelset and pantograph. These two components are one of the most important components of the

entire metro vehicle. The performance of the subway directly determines whether the subway vehicle can operate safely and stably.

In this paper, the Markov chain is used to make state maintenance decisions for both. First, it is assumed that the faults and maintenance processes of the pantograph and wheelset are independent of each other and are not related to each other [12]. Since the storage environment of the equipment is relatively fixed, the judgment of the state is mainly based on the number of repairs and the time of use, and the experience is judged mainly based on the number of repairs. The maintenance decision model is shown in Fig. 4.

According to the above flow chart, we can see that in the process of equipment maintenance decision-making, we first set the maintenance decision-making between the equipment does not affect each other, then divide the key components of the maintenance decision-making, and make maintenance decision for each component, and get the results of maintenance decision-making and the corresponding maintenance time. A minimum maintenance time is further determined, and the corresponding time difference is obtained by comparing the difference values. If the time difference is greater than the nominal value, it means that there is no need for maintenance. In the next planned maintenance, the maintenance of parts with longer maintenance time is neglected, and the parts are repaired in stages. If the time difference between

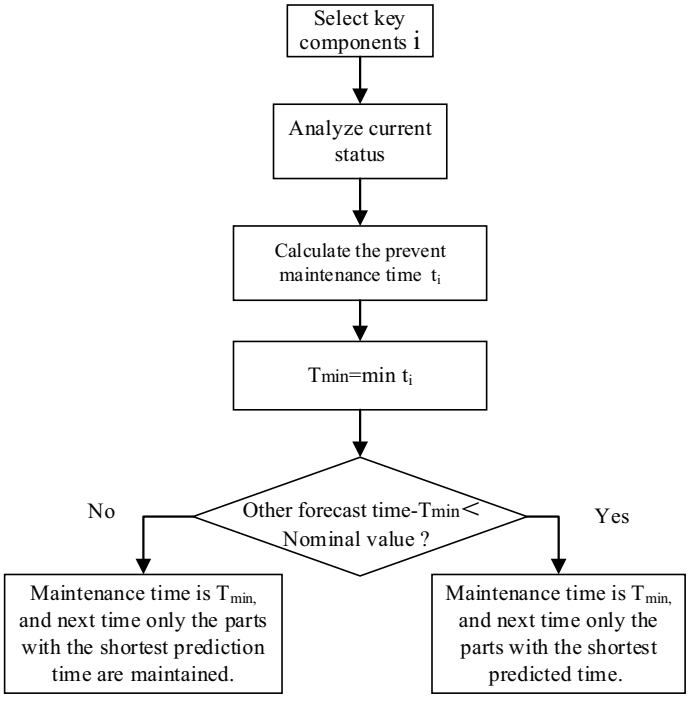


Fig. 4 State decision model based on Markov chain

the two parts is small, the maintenance of the two parts is combined to reduce the impact of too many maintenance times on the reliability of equipment.

According to the above analysis, it can be seen that the main factors affecting the state of each component of the equipment are the number of maintenance times n and the after maintenance time t. While the storage environment of the equipment and the technical level of the maintenance personnel can be neglected.

Number the key parts of the Metro as $\delta = 1, 2, \ldots, N$, and set:

(1) Obtain the state judgment data function: $f = f(t, n)$
(2) Transfer matrix function: $R = R(h)$
(3) Judging data state function: $H = h(f)$
(4) Transfer probability function: $A_{ij} = M_{ij} = (I - R)_{ij}^{-1}$
(5) Maintenance time function of single component: $T = (j > i)A_{ij}\ T_j$.

Then the component delta fault prediction time is:

$$T_\delta = \sum\nolimits_{j \geq i} \left\{ [I - R(h(f(t, n)))]_{ij}^{-1} T_j \right\} \tag{5}$$

Maintenance decision time function:

$$T_{min} = \min(T\delta) = \min(T_1, T_2, \ldots, T_n) \tag{6}$$

Maintenance time difference function: $\Delta T_\delta = T_\delta - T_{min}$.

Based on the Markov chain state maintenance decision process, the state data of the equipment is first obtained according to the number of maintenance and the use time after maintenance. The state level of the equipment at the moment is determined according to the state data, the state transition matrix A is listed, and the state transition is calculated. Probability, further calculate the preventive maintenance time of a single equipment component, and solve the T_{min} and the difference between each repair time and the minimum value. Then compare with the set time difference nominal value to obtain the corresponding maintenance decision.

4 Case Analysis of State Maintenance Decision Model

This paper analyzes the annual actual maintenance data of wheelsets and pantographs in a domestic subway company in 2017. The wheelset and pantograph are divided into four states: good, normal, degraded and intended to be faulty, which are denoted as B_1, B_2, B_3, B_4 and D_1, D_2, D3, D4 and fault status K.

The evaluation of the condition of the components is mainly based on the number n of repairs and the time t used. According to the historical fault data of the component and the maintenance detection information, through the Matlab data processing and fitting, the correlation function expression of the influence factor of the wheelset is obtained:

Table 1 Wheel pair status judgment table

f(t, n)	(0, 0.5)	(0.5, 1.2)	(1.2, 1.8)	(1.8, 2.6)	(2.6, ∞)
State	B_1	B_2	B_3	B_4	K

Table 2 Pantograph pair status judgment table

f(t, n)	(0, 0.2)	(0.2, 0.4)	(0.4, 1)	(1, 1.8)	(1.8, ∞)
State	D_1	D_2	D_3	D_4	K

$$\mathrm{f(t,\,n)}\left[1 - exp\left(-\frac{t}{928}\right)\right] \times 8 + \frac{n}{36} \tag{7}$$

The relevant function expression of the influence factor of the pantograph is:

$$\mathrm{f(t,\,n)}\left[1 - exp\left(-\frac{t}{1869}\right)\right] \times 5 + \frac{n}{20} \tag{8}$$

The state judgment of the wheel pair is shown in Table 1.

The state judgment of the pantograph pair is shown in Table 2.

If the number of repairs is $n = 0$, the service time after repair is $t = 200$d, the number of repairs of the pantograph is $n = 0$, and the time after repair is $t = 80$d, that is, the state of the wheelset is B_1, and the power is received. The state of the bow is C_2, and the state prediction maintenance time of the two components is calculated separately. Taking the wheel pair as an example, the state of each component stays is shown in Table 3.

According to the historical data analysis, the transition probability of the wheel is shown in Table 4.

List the state transition matrix A:

Table 3 Wheels for each state stay schedule

State	B_1	B_2	B_3	B_4
Time/d	175	100	65	36

Table 4 Wheelset transfer probability table

State	$B_1 \rightarrow B_1$	$B_1 \rightarrow B_2$	$B_1 \rightarrow B_3$	$B_1 \rightarrow B_4$	$B_1 \rightarrow D$
Probability	0.68	0.52	0.2	0.09	0.01

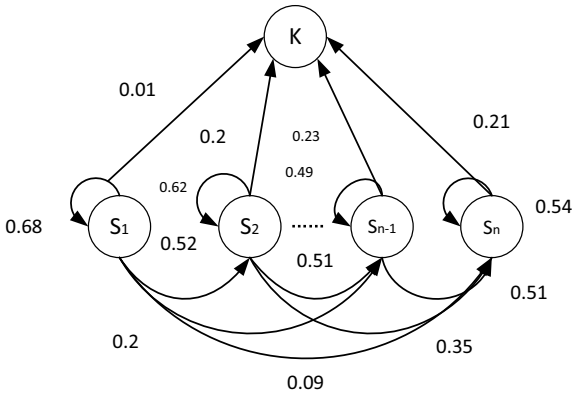


Fig. 5 Wheelset Markov state transition diagram

$$\begin{bmatrix} 0.68 & 0.52 & 0.2 & 0.09 & 0.01 \\ 0 & 0.62 & 0.51 & 0.35 & 0.2 \\ 0 & 0 & 0.49 & 0.51 & 0.23 \\ 0 & 0 & 0 & 0.54 & 0.21 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The Markov chain state transition model is shown in Fig. 5.

Calculate the predicted repair time when the wheel pair is in this state:

$$T = 2.236 * 175 + 1.582 * 100 + 1.353 * 65 + 4.324 * 36 \approx 790d$$

Similarly, the pantograph predicted maintenance time $T' \approx 310d$.

The predicted maintenance time difference between the two components is 480d, and the time difference between the two is long, so in the next preventive maintenance, only the maintenance of the pantograph can be considered first, without the need for preventive maintenance of the wheelset.

If the number of the wheeled maintenance is $n = 1$, the service time after repair is $t = 150d$, the number of the pantograph maintenance is $n = 1$, and the service time after maintenance is $t = 150d$, that is, the state of the two is at B_3 respectively. With the C_3 state, the predicted maintenance time of the two can be obtained by Markov chain prediction as $T = 350d$ and $T' = 323d$, and the time difference is 27d. The preventive maintenance time of the two is small, so the next preventive maintenance should be performed on both components at the same time to reduce the impact on the reliability of the equipment itself due to excessive maintenance.

5 Conclusion

Based on the analysis of the characteristics of Markov chain state transition, this paper find that its characteristics are in line with the transfer of key equipment maintenance status of rail transit. Therefore, this paper proposes a dynamic state maintenance decision model based on Markov chain. The traditional maintenance guarantee theory of the "planned repair" and "state repair" of the metro has realized the optimal maintenance decision-making plan. Through the research on the key factors such as the number of maintenance and the time after maintenance, the optimal maintenance time interval of the equipment was pre-judgated, which reduced the reliability of the equipment due to the frequent maintenance. The key equipments studied in this paper are few, the follow-up research can be established in a higher dimension, and combined with big data and other technologies to establish a perfect maintenance support platform.

References

1. Zhang Z (2015) Study on the new anti-ship missile optimization decision based on condition maintenance. Naval Aeronautical and Engineering University, Yantai, pp 1–10
2. Cassady CR, Murdock WP, Pohl EA (2001) Selective maintenance for support equipment involving multiple maintenance actions. Eur J Oper Res 129(2):250–259
3. Conroy GV, Black WJ, O'Hare GMP (1989) Fault diagnosis and on condition maintenance of a combined heat and power unit. In: Proceeding of IEEE colloquium on expert systems for fault diagnosis in engineering applications. IEEE, London, pp 3–6
4. Tomita N, Beldica C (1988) On-condition maintenance of machinery based on vibration measurement and frequency analysis. Buletinul Institutului Polytechnic Backrest, Serial Mechanical
5. Zhang X, Zeng J (2015) Opportunity maintenance optimization of two-component system based on state space segmentation. Syst Eng Theor Pract 6:10–17
6. Dong Y, Wang X, Gu Y (2008) Risk-based maintenance decision and its application in power station equipment. Mod Electr Power 25(2):51–56
7. Xin Y, Yue D, Gao S et al (2011) Research on state maintenance decision model based on non-ideal maintenance. Ship Electron Eng 31(10):130–133
8. Zhang G, Wang Y, Fan W et al (2015) Research on dynamic decision model of equipment state maintenance monitoring interval using Markov chain. J Sichuan Ordnance Ind 36(4):81–84
9. Su C, Zhou X (2012) Wind turbine state maintenance optimization based on semi-Markov decision process. J Mech Eng 48(2):44–49
10. Yanxue Xu, Jin An (2018) Markov-based equipment state maintenance decision model. J Gun Launch Control 39(3):2–7
11. Department of Mathematics, Tongji University (2007) Engineering mathematics linear algebra, 5th edn. Higher Education Press, Beijing
12. Gan M, Kang J, Gao Q (2005) Military equipment maintenance engineering. National Defense Industry Press, Beijing, pp 210–221

Research on the Choice of Shared Car Travel Behavior Based on Medium Commuting Distance

Yongneng Xu, Xiaotian Wang and Zhou He

Abstract With the rapid development of urbanization level, the choice of urban residents to travel under medium commuting distance has a great impact on urban traffic congestion and environmental pollution. Shared car travel modes can alleviate traffic pressure. This paper takes the travel choice behavior of residents under the medium commuting distance as the research object. Through the t-test to the age, the number of private cars, whether there is car rental experience, travel expenses and travel time as the influencing factors. Multivariate Logit selection probability model of the traveler's choice of private car, shared car, taxi and net car is established. The SPSS software is used to calibrate and test the model parameters by using the maximum likelihood estimation method to obtain the travel mode selection probability, which provides a reference for the development of shared car mode.

Keywords Medium commute distance · Shared car · Multivariate Logit model · SPSS software

1 Introduction

With the rapid advancement of urbanization and the rapid growth of urban population, the number of private cars in residents has gradually increased. According to the statistical results, the average annual growth rate of private car ownership in first-tier cities in China has reached 20%. The proportion of private car travel in urban traffic is also increasing, which further exacerbates the problem of urban road traffic congestion. According to the study of the Gaode map, compared with 2015, 80% of the 45 major cities surveyed in 2016 have an increased delay congestion index [1].

Y. Xu · X. Wang (✉) · Z. He
Nanjing University of Science and Technology, Transport Engineering, Nanjing, China
e-mail: 1350495938@qq.com

Y. Xu
e-mail: x780906yn@163.com

Z. He
e-mail: 1572484107@qq.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_109

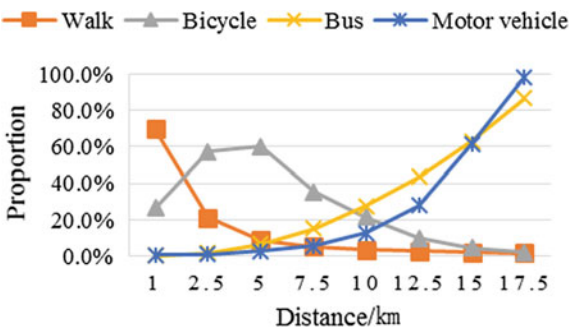


Fig. 1 Suzhou travel distance curve

Commuting travel is an important way for urban residents to travel, with a single travel path, stable time, and tidal traffic. Especially in the morning and evening commute peak hours, the road congestion problem is even more serious. Medium commute distance refers to 10–60 km or driving private car time is 30–60 min [2]. It can be seen from the travel distance curve of Suzhou City in Fig. 1. When the travel distance is less than 10 km, there are more residents who choose to walk and bicycle. When the travel distance is greater than 10 km, the residents are more inclined to choose the bus and motor vehicle travel modes [3].

Public transportation is favored by the government and travellers with high carrying capacity, environmental friendliness, and convenient and punctuality. However, public transportation passes through low-density suburbs, and the station density is not high, and there is a certain distance from the work place and home. If the commuter drives a private car to the destination every day, the transportation cost is relatively high, and on the other hand, the parking space is tight. As a new type of transportation, shared vehicles are distributed in the vicinity of public transportation hubs and commercial centers. Shared cars provide rental services at low cost, point-to-point to fill the gap in public transportation, and solve the traffic demand for the last commute travel distance from the station to the destination. Take the new energy vehicle time-sharing EVCARD as an example. There are about 6000 outlets in Shanghai, mainly in transportation hubs, business districts, large communities and hotels. 80% of the outlets are more than 1 km away from the nearest subway station. Therefore, sharing cars can make up for the blind spots of urban public transportation [4]. The shared car travel mode not only relieves urban congestion, energy and environmental problems, but also makes commuters more economical, fast and comfortable.

Shared cars can be divided into 2 categories according to the pick-and-ride mode: round-trip shared cars, each car has a fixed return point, return to the same parking point after picking up at the parking point; one-way shared car, no fixed. For the parking spot, the owner can choose another parking spot in the city to change the car. Shared car charging methods use time-sharing, including time cost, distance cost, and fuel cost. The use of shared cars is mainly divided into three steps: after registering and perfecting personal information and paying membership fees, become a member of the shared car organization, use mobile phone software to reserve nearby

vehicles; arrive at the outlet to pick up the car; return the vehicle to the pre-selected outlet settlement [5].

Dai Qiujie [6] mainly analyzes the factors affecting consumers' choice to join the shared automobile organization from four aspects: family economic conditions, related expenses, car demand and personal characteristics. Kim [7] investigated the substitutability of shared cars for the purchase of private cars, from car use, purchase costs, maintenance costs, operating costs, time costs, and walking time. Liu Wei [8] built an evolutionary game model for the game objects by travellers, governments, travellers and car operators, and obtained government policies and operators' large-scale operations to help guide travellers to choose to share cars. Consumers tend to choose shared car services within 1 km. For personal commuter users who use shared cars, it is usually more convenient to pick up a car or to go to work in areas where rail transit buses have not arrived [9]. The area where the commuters live is mainly residential development, and the natural environment is better, generally far from the public transportation station [10].

Han Xiaoyu et al. [11] applied the multivariate Logit model to predict the orbital connection mode, and calibrated the model parameters by SPSS to obtain the natural logarithmic model of the sharing rate of walking, bicycle/electric vehicles and public bicycles relative to buses. Guo Yuxin [12] analyzed the characteristics of Beijing rail transit group and rail transit chain, and established a three-tier rail transit Logit model based on utility maximization theory. The parameter calibration is solved by SPSS software, and the sharing rate of each rail transit link of each station is obtained. Jing Peng [13] pointed out that the intrinsic reasons for the formation of individual preferences in the selection process and some factors that cannot be directly observed are used as "black boxes" in traditional discrete analysis and cannot be explained. In this paper, the theory of planned behavior is used as a framework to integrate various psychological factors that have an impact on the choice of travel modes. The hybrid selection model is used to study the role of psychological influence factors in the choice of travel modes in cities. Zhu Shunying [14] fuzzed the travel time, ignoring the impact of travel expenses, and established a Logit model for travel mode with fuzzy characteristic variables. Yang Liya [15] established two resident NL models with the departure time at the lower level and the travel mode at the lower level.

2 Model Theory

The non-aggregate model of the above-mentioned travel mode prediction considers the random uncertainty of the total utility, and the characteristic variable value in the utility fixed item is a real number, that is, the traveler's expected value of the travel mode selection, such as the travel time and the travel cost, is a certain value. The value, plus the error term. Travelers as decision makers will compare the attributes of different modes of transportation and choose the mode of travel based on the principle of utility maximization. This is the model basis for Logit and a prerequisite for all non-aggregate models to be obeyed. Assume that U_{in} in is the utility of the traveler

n to choose the travel plan i, and A_n is the set of alternative travel modes. According to the discrete selection model of the random utility theory, U_{in} in is expressed as follows.

$$U_{in} = V_{in} + \varepsilon_{in} \tag{1}$$

V_i represents the portion of the scheme i utility that can be observed, and ε_{in} represents the random portion of the scheme i utility.

$$V_{in} = \beta_1 X_{in1} + \beta_2 X_{in2} + \cdots + \beta_k X_{ink} = \sum_{k=1}^{K} \beta_k X_{ink} = \sum_{i \in A_n} \beta' X_{in} \tag{2}$$

X_{in_1} indicates that for different attributes in the traveler n scheme i, β_1 is the weight of the usage scheme attribute.

The condition that the traveler n selects the scheme i from the set A_n is:

$$U_{in} > U_{jn}, \mathrm{j} \in A_n \tag{3}$$

The multivariate Logit model is one of the commonly used non-aggregate models. It is suitable for distribution selection when utility is maximized and is widely used in transportation mode selection. The probability formula for individual n to choose scheme i is:

$$P_{in} = \frac{\exp(V_{in})}{\sum_{j \in A_n} \exp\left(V_{jn}\right)} = \frac{1}{\sum_{j \in A_n} \exp\left(V_{in} - V_{jn}\right)}, \mathrm{j} \in A_n \tag{4}$$

$$0 \le P_{in} \le 1, \sum_{i \in A_n} P_{in} = 1 \tag{5}$$

The Logit model follows four assumptions:

(1) All travellers have the same form of utility function.
(2) The parameters of the utility function do not change with the traveler and are relatively stable.
(3) The random part of the utility function of each traveler for different limbs is independent of each other and obeys the gumbel distribution of the same variance.
(4) The traveler chooses the largest selection limb of the utility function.

3 Model Establishment

The multivariate Logit regression model is established to determine the traffic mode-bearing rate of the commute from the place of residence to the public transportation site.

A. *Variable selection is explained*

The explanatory variable is a collection of all modes of transportation. Usually, in the case of inconvenient public transportation, the commuting modes suitable for medium travel distance (10–60 km) are: private cars, shared cars, taxis, and cars.

The explanatory variables mainly select the important factors that affect the choice of travellers: the characteristics of the traveler (age, gender, personal income, whether they own a private car, whether there is a car rental experience), travel characteristics (travel distance, travel time, travel expenses). Since some of the influencing factors in the questionnaire cannot be quantified, each influencing factor will be assigned.

The travel time of the private car is the ratio of the travel distance to its speed; the travel time of the shared car is the walk pickup time + driving time + walking return time; the taxi travel time is the taxi waiting time + the car time.

The travel expenses of private cars are fuel consumption + parking lot parking fees; the travel expenses of time-shared (shared cars) are shared car charges; the taxi fare is about taxi fare.

The following eight factors affect the factors that commuter travellers choose transportation mode. According to the actual situation and correlation analysis, factors that are significantly related to the choice of transportation mode can be selected as explanatory variables (Table 1).

B. *Share Rate Model*

Equation (4) is the basic form of the multivariate Logit model, and the Logit model of the SPSS software output is its deformed form, as shown in the following figure.

$$\mathrm{Ln}\left(\frac{P(y=j|x)}{P(y=J|x)}\right) = a_j + \sum_{k=1}^{K} b_{jk}x_k = V_j \tag{6}$$

$$\sum_{j=1}^{J} P(y=j|x) = 1 \tag{7}$$

4 Model Application

This paper selects travellers whose commuting distance is 10–60 km. The survey content is mainly the three characteristics of traveler's choice of will, personal attributes and travel characteristics, and finally 1000 questionnaires. According to

Table 1 Description of factors affecting travel selection behavior

Characteristic variable			Description
Individual attribute	Gender	x_1	Male $x_1 = 1$, female $x_2 = 2$
	Age stage	x_2	Segmentation based on the age range of the traveler $x_2 = 1, 2, 3, \ldots$
	Personal income	x_3	Segmentation based on the income interval of the traveler $x_2 = 1, 2, 3, \ldots$
	Whether you own a car	x_4	There is a car $x_4 = 1$, no car $x_4 = 0$
	Whether you have a car rental experience	x_5	Yes $x_5 = 1$, no $x_5 = 0$
Travel characteristics	Travel distance	x_6	Assignment based on actual distance
	Travel time	x_7	Assignment based on actual time
	Travel expenses	x_8	Assignment based on actual cost

the questionnaire, SPSS data processing software was used to analyze the influencing factors.

Processing extreme data and checking the multicollinearity of the data, since the maximum VIF is 2.17, much less than 10, there is no need to worry about multicollinearity.

A. *Interpreting variable assignments*

A t-test is performed on the independent variables, and the age, the private car, the travel expenses, the rental experience and the time are selected as explanatory variables, and a multi-dimensional Logit model based on the medium commuting distance-travel selection behavior is established. As shown in Table 2.

Table 2 Travel traffic mode variable assignment

	Variable		Private car	Shared car	Taxi/net car
Individual attribute	Age	x_1	18–30, 31–40, 41–50, over 50 years old		
	Have a car	x_2	There is one = 1, there are two = 2, no = 3		
	Car rental experience	x_3	Renting experience = 1, no car rental experience = 2		
Travel characteristics	Travel expenses	x_4	Assignment based on actual cost		
	Travel time	x_5	Assignment based on actual time		

Table 3 Calibration of the share rate parameter (taking the shared car as an example)

Transportation		B	Wald	Degree of freedom	Significant
Shared car	Intercept	18.647	0.000	1	0.997
	Travel expenses	0.163	9.790	1	0.002
	Time	−0.028	1.389	1	0.239
	[age = 1]	13.181	0.000	1	0.998
	[age = 2]	15.139	0.000	1	0.998
	[age = 3]	16.840	0.000	1	0.997
	[age = 4]	0^h	–	0	–
	[Car ownership = 1]	−0.479	0.118	1	0.731
	[Car ownership = 2]	0.353	–	1	–
	[Car ownership = 3]	0^h	–	0	–
	[Car rental experience = 1]	0.218	0.040	1	0.841
	[Car rental experience = 2]	0^h	–	0	–

B. *Parameter calibration*

Taking the taxi/net car as the reference category, the model parameters are calibrated by the multi-Logit analysis in the SPSS software (taking the shared car as an example), as shown in Table 3.

Therefore, the natural logarithmic model of the sharing rate of the private car and the shared car relative to the taxi/net car can be obtained separately:

$$\begin{aligned}\text{Logit}\frac{P_1}{P_3} = \ln\left(\frac{P(y=1|x)}{P(y=3|x)}\right) &= 0.020 * x_4 \\ &- 0.021 * x_5 + 0.052 * (x_1 = 1) + 2.546 * (x_1 = 2) \\ &+ 1.798 * (x_1 = 3) + 3.424 * (x_2 = 1) \\ &+ 19.168 * (x_2 = 2) - 1.260 * (x_3 = 1) - 2.046 \end{aligned} \tag{8}$$

$$\begin{aligned}\text{Logit}\frac{P_2}{P_3} = \ln\left(\frac{P(y=2|x)}{P(y=3|x)}\right) &= 0.163 * x_4 - 0.028 * x_5 \\ &+ 13.181 * (x_1 = 1) + 15.139 * (x_1 = 2) \\ &+ 16.840 * (x_1 = 3) - 4.479 * (x_2 = 1) \\ &+ 0.353 * (x_2 = 2) + 0.218 * (x_3 = 1) - 18.647 \end{aligned} \tag{9}$$

Table 4 Model fitting information

Model	Model fitting condition		Likelihood ratio test	
	−2 log likelihood	Bangla	Degree of freedom	Significant
Intercept only	146.963	–	–	–
finally	71.908	75.054	16	0.000

Table 5 Pseudo R square

Model	Pseudo R side
Cox-snell	0.628
Negorco	0.730
McFadden	0.503

C. *Model test*

The SPSS software can verify the reliability of the parameter calibration results:

(1) Model fitting information. Table 4 shows the model fitting information for the output. The final significance is 0.000, and the significance level is 0.05, indicating that the linear relationship between the explanatory variables, whether there is a private car, travel expenses, whether there is a car rental experience and time, and the taxi/network car sharing rate is significant. The model variables are selected correctly.

(2) Goodness of fit indicator. Table 5 shows the goodness of fit indicators. The pseudo R square value is between 0 and 1, and the better the fitting effect of the model. The McFarden pseudo R-side is generally ideal between 0.3 and 0.5 [11]. Therefore, the model McFadden pseudo R square is 0.503, and the fitting effect is better as shown in Table 3.

(3) Likelihood ratio test. The likelihood ratio test can obtain the variation of the likelihood ratio chi-square value and the −2 log-likelihood value of the simplified model after the model introduces or extracts a certain variable. According to the likelihood ratio test, when the significance level is 0.05, the cost, age, and the probability P value of the chi-square test with the automobile condition are 0.000, 0.048, and 0.000, respectively, and the assumption that the regression coefficient is 0 is rejected. Age and ownership of the car have a greater impact on the linear relationship of the model. Time and car rental experience have no obvious influence on the linear relationship of the model. When choosing the explanatory variables, the time and rental experience can be used as the explanatory variables of the non-forced entry model to improve the fitting effect of the model.

D. *Model prediction*

The prediction results of the model are shown in Table 6. Usually, when the hit rate of the model reaches 80% or more, the model can be considered to have better results

Table 6 Model prediction classification

	Prediction			
Measured	Private car	Shared car	Taxi/net car	Correct percentage (%)
Private car	434	13	53	86.8
Shared car	26	92	13	70.2
Taxi/net car	39	1	329	89.2
Overall percentage (%)	49.9	10.6	39.5	86.0

[8]. It can be seen from the table that the correct percentages of private car, shared car and taxi/net car are 86.8, 70.2, 89.2%, and the overall hit rate is 86.0%. Therefore, it can be stated that the calculation result of the model is acceptable.

5 Conclusion

In this paper, we use the multivariate Logit model to analyze the traffic behavior choice behavior of middle commuter distance travellers, and finally get the traffic mode selection model. The cost, age and ownership of the car have a significant impact on the medium commuting distance traveler's traffic mode selection behavior.

In view of the good environmental, economic and social benefits of shared vehicles, in the future development, we must first popularize the awareness of sharing cars, so that middle-class commuters of different ages can understand the use of shared cars. Second, the government can Enterprises and individuals using shared cars provide subsidies to guide more consumers to use shared cars from the cost; in the end, the number of private cars has a greater impact on the use of shared cars by medium commuters, and the government should encourage everyone to reduce private cars. The number of purchases has shifted from using private cars to using shared cars.

Although this paper has certain reference to the development of shared vehicles, the model is too small and ignores the multi-logit model. The irrelevant choice independence hypothesis (IIA) characteristics and randomness of preferences, residents' behavior habits, driver's responsibility The impact of factors such as division, vehicle comfort and accessibility on the model requires further study.

References

1. Kearney "互联网 + 智能交通"市场进入发展黄金期. 上海汽车, 2018(1):61–62
2. 石飞, 陆振波. 出行距离分布模型及参数研究. 交通运输工程学报, 2008, 8(2):110–115
3. 陈征, 刘英舜, 石飞. 城市居民出行方式距离曲线的研究与应用. 交通运输研究, 2006(8):157–160

4. 荣萍, 罗 祎辰. 引领汽车产业新趋势——EVCARD电动汽车分时租赁助力“智慧交通”建设. 中国高新区, 2016(17):146–150
5. 程苑. 汽车共享下的城市交通出行方式博弈研究. 哈尔滨工业大学, 2015
6. 代秋杰.基于消费者选择行为的汽车共享政策分析. 长春理工大学学报(社会科学版), 2012, 25(11):106–107 + 120
7. Kim J, Rasouli S, Timmermans H (2017) Satisfaction and uncertainty in car-sharing decisions: an integration of hybrid choice and random regret-based models. Transp Res Part A 95:13–33
8. 刘珺. 汽车共享出行模式选择的演化博弈研究. 南京大学, 2017
9. 兰静, 诸大建. 可持续交通消费的接受和使用行为研究——基于上海市汽车共享的调查. 中国人口资源与环境, 2016, 26(11):98–105
10. 张明, 刘菁. 适合中国城市特征的TOD规划设计原则. 城市规划学刊, 2007(1):91–96
11. 韩晓玉, 朱从坤, 何承韡,等. Logit模型在轨道交通接驳方式预测中的应用. 交通科技与经济, 2016, 18(1):25–29
12. 郭彧鑫. 基于出行链的轨道交通衔接方式研究[D]. 北京建筑大学, 2015
13. 景鹏, 隽志才, 查奇芬. 考虑心理潜变量的出行方式选择行为模型. 中国公路学报, 2014, 27(11):84–92
14. 朱顺应, 邓爽, 王红,等. 具有模糊特性变量的出行方式预测Logit模型. 交通运输工程学报, 2013, 13(3):71–78
15. 杨励雅, 邵春福, HAGHANI A. 出行方式与出发时间联合选择的分层Logit模型. 交通运输工程学报, 2012, 12(2):76–83

Study on Passenger Flow Characteristics and Classification Method of Rail Transit Stations Based on AFC Data—A Case Study of Ancient District of Suzhou

Peipei Peng and Daixiao Zou

Abstract The passenger flow characteristics of rail transit stations are important data bases for urban rail transit development planning, rail transit station connection strategy designation, underground space planning of rail transit stations and surrounding land development. Traditionally, the characteristics of rail passenger flow are acquired by manual survey method, and the continuous passenger flow characteristics cannot be obtained. In this paper, AFC data is used to analyze the passenger flow characteristics of rail stations, and the classification method of rail stations is studied according to the different characteristics of passenger flow in and out of different stations. Finally, the passenger flow and station classification of rail stations in the ancient urban area of Suzhou is taken as an example.

Keywords Rail transit stations · AFC data · Passenger flow analysis · Classification method · Ancient Urban Area of Suzhou

1 Introduction

A fundamental basis for rail transit construction is to grasp how it operates. As rail transit lines and networks have complicated functions and levels, their overall efficiency and level differences should be part of corresponding layout and planning [1]; in this process, core links are to identify passenger flow demands, install station facilities and realize station connection [2].

Rail transit station (RTS)-influenced areas are the crucial key to the transit-oriented development mode designed to guide urban development [3]. RTSs vary in passenger flow characteristics. Before a city initiates relevant studies, such as future urban rail

P. Peng (✉)
Transportation College of Southeast University, Nanjing 210096, Jiangsu, China
e-mail: 464980255@qq.com

Suzhou Rail Transit Company, Suzhou 215004, Jiangsu, China

D. Zou
Suzhou Plan and Design Company, Suzhou 215004, Jiangsu, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_110

transit development planning, RTS connection strategy formation, RTS accessibility research, land development guidance and land use, a science-based and proper classification of RTS-influenced areas is necessary [4].

Existing studies take manual survey as a common way to analyze the characteristics of rail passenger flows. However, this method has two drawbacks: manpower-consuming survey process and low-precision survey result. In contrast, using data based on transit AFC (automatic fare collection system, where a passenger swipes a card to enter/exit a station) means higher precision and more reliable analysis results. This paper, therefore, relies on transit swiping data to, probe into how stations are different in passenger flow characteristics, and further study how to categorize these stations, in an effort to offer data support for urban rail transit planning.

2 An Analysis of Rail Transit Passenger Flow Collector-Distributor Point Classification

Analyzing the RTS characteristics is a major task of point-layer features. In doing so, the aims are twofold. Firstly, swiping data recorded when passengers enter/exit a RTS can be garnered to analyze indicators relating to such passenger flows, like temporal changes, spatial changes and entering/exiting numbers [5]; secondly, depending on the usage nature of different kinds of land, RTSs can be classified into specific types accordingly. Furthermore, understanding how various stations interact with the surrounding transport environment [1] is useful for site selection, operation and management with respect to RTS.

Differences in nature of land usage mean differences in type of passenger flow collector-distributor points, so grouping these points and studying their distinctive features should be necessary given the special demand for rail transit [6]. Table 1 sets out the classification results. Main collector-distributor points are located in residential areas, administration areas, business centers, cultural centers, healthcare centers and scenic spots. Capturing the characteristics of passenger flows within these points is a precondition for understanding rail transit's operation situation.

3 Study of Rail Transit Station Classification

3.1 Analysis Concept

Passenger flows entering and exiting stations feature temporal changes. The study starts with clarifying functions of various typical RTSs based on features of such changes, before identifying functions of other stations by category. Then, in light of the characteristics of passenger flows at various types of stations, the study

Table 1 Classification of rail transit passenger flow collector-distributor points

Classification of passenger flow collector-distributor points	Scope of collector-distributor points	Characteristics of collected/distributed passenger flows
Transport hubs	Railway stations, bus stations, airports	Hub transfer and passenger flow connection
Residential areas	Large residential areas	Resident commuting and flexible passenger flows
Industrial areas	Industrial parks, special industry areas	Worker commuting and business travel
Administration centers	Main government offices	Worker commuting and official travel
Shopping centers	Large malls and shopping centers	Shopping and consuming passenger flows
Cultural centers	Cultural central areas	Study and entertainment passenger flows
Healthcare centers	Large hospitals	Doctor-seeing passenger flows
Sports centers	Large sports stadiums	Exercising and fitness passenger flows
Scenic spots	Main tourism attractions	Tourism and leisure passenger flows

aims to propose suggestions for preparing integrated transport facilities, and provide reference for integrated planning and optimization of rail transit lines in the future.

There are already studies centered on the integration among rail transit lines in Suzhou. RTSs are classified into origin/destination stations, hub oriented stations, shopping oriented stations, office oriented stations, dwelling oriented stations, tourism oriented stations, and mixed shopping + office/shopping + dwelling oriented stations. According to the use of land surrounding current stations, some typical ones are preliminarily selected [7], as shown in Table 2.

3.2 Classification Standards

By referring to the existing research result in relation to station classification, and taking into account characteristics of passenger flows at different RTSs, the paper summarizes standards to distinguish one type of station from another. As Table 3 shows, there are mainly four kinds of stations, namely, dwelling-oriented, office-oriented, shopping-oriented and tourism-oriented. For the first three kinds, the judging standard is the distribution of time to enter/exit a station; and for the last one,

Table 2 Suzhou RTS classification

Station classification	Station characteristics	Typical stations
Origin/destination stations	As the starting/ending point of rail transit lines, origin/destination stations usually sit at the edge or the periphery of the downtown area, and function as transfer stations to serve a wide range in the surrounding area	Mudu, Zhongnan Jie, Qihe Lu, Sangtiandao, Longdaobang, Tongli
Hub oriented stations	Main passenger transport stations connecting a city with the outside, like airports, railway stations and long-distance bus stations, together with rail transit, form a large integrated transport transfer hub system of the city	Suzhou Railway Station, Suzhou North Railway Station
Shopping oriented stations	Shopping oriented stations are close to city-level and district-level public shopping centers. Given the number of large public buildings, passenger flows are huge. These stations can reach the whole urban area and even the more peripheral towns	Lindun Lu, Times Square, Shi Lu, Chayuanchang
Office oriented stations	Office oriented stations are close to land mainly used for business and administrative office functions	Xinghu Jie, Xinghai Square
Dwelling oriented stations	Dwelling oriented stations are close to large communities where land mainly serves residential purposes	Donghuan Lu, Panli Lu
Tourism oriented stations	Tourism oriented stations are close to large scenic spots and mainly serve tourism purposes	Suzhou Amusement Land, Shantang Jie
Mixed functions oriented stations	Mixed functions oriented stations adjoin areas that feature mixed types of land use and serve two or more functions, all with relatively large buildings	

Table 3 Standards to identify station type

Station type	Identification standard (Time for passengers to enter/exit a station)
Dwelling-oriented	On weekdays, entering rate at 8:00 a.m. is larger than 18%, and exiting rate at 8:00 a.m. is less than 12%; while entering rate at 5:00 p.m. is less than 10%, and exiting rate at 5:00 p.m. is larger than 12%
Office-oriented	On weekdays, entering rate at 8:00 a.m. is less than 10%, and exiting rate at 8:00 a.m. is larger than 20%; while entering rate at 5:00 p.m. is larger than 14%, and exiting rate at 5:00 p.m. is less than 10%
Shopping-oriented	On weekdays, entering rate at 8:00 a.m. is less than 8%, exiting rate at 8:00 a.m. is less than 15%, and entering rate at 8:00 p.m.–9:00 p.m. is larger than 5%; on weekends, over 3 h' exiting rate is larger than 7% during 10:00 a.m.–3:00 p.m., and over 4 h' entering rate is larger than 7% during 3:00 p.m.–9:00 p.m.
Tourism-oriented	Entering rate on weekdays $\leq$ entering rate on weekends $\leq$ entering rate during holidays

the identification principle is that the entering rate on weekdays is less than that of weekends, which is also less than the entering rate during holidays.

4 Case Study—Ancient Urban Area of Suzhou

4.1 RTS Distribution in Ancient Urban Area of Suzhou

Ancient Urban Area of Suzhou now has two rail transit lines, including Line 1 and Line 4 that include eight stations in total, as sketched in Fig. 1.

Under the said standards to identify the RTS type, the paper incorporates Suzhou's rail transit swiping data, in order to classify existing RTSs in Ancient Urban Area of Suzhou. It is found RTSs inside the Area fall into four types, namely, dwelling-, shopping-, tourism- and office-oriented stations. Table 4 matches each station name to the corresponding station type. Clearly, shopping- and office-oriented stations dominate throughout the Area.

4.2 Characteristics of Changes in Passenger Flows Entering/Exiting Different Types of Stations

(1) **Dwelling-oriented Stations**

Table 5 and Fig. 2 illustrate changes in passenger flows entering/exiting Nanmen and Xiangmen stations. Specifically, during morning rush hours, most passengers are entering the stations, making up over 20% of the total. Entering passenger flows

Fig. 1 Map of RTS distribution in Ancient Urban Area of Suzhou

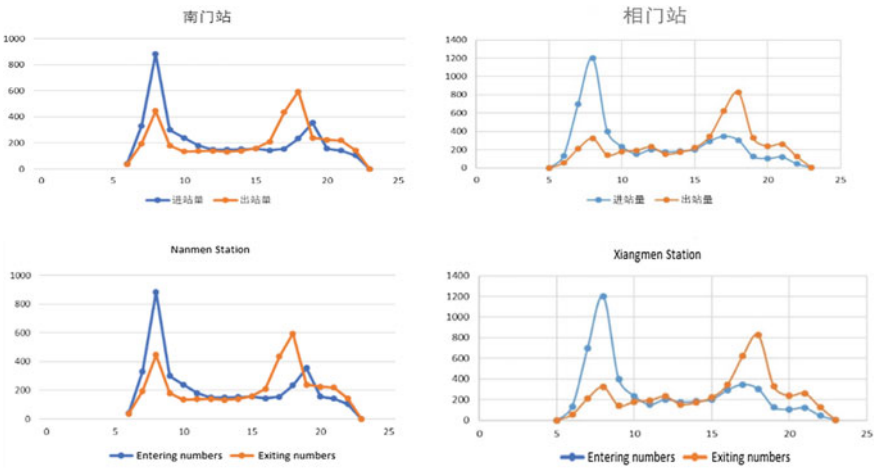
Table 4 Table of function classification and identification of RTSs in Ancient Urban Area of Suzhou

Station name	Station type
Yangyu Xiang	Office-oriented
Leqiao	Shopping-oriented
Lindun Lu	Shopping-oriented
Xiangmen	Dwelling-oriented
Beisita	Tourism-oriented
Chayuanchang	Shopping-oriented
Sanyuanfang	Office-oriented
Nanmen	Dwelling-oriented

are over twice that of exiting passenger flows; during evening rush hours, most passengers are exiting the stations, taking up over 15% of the total. And exiting passenger flows are over 1.5 times larger than entering passenger flows. During the remaining non-rush hours (from 10:00 a.m. to 3:00 p.m., the same below), the proportion of entering/exiting passenger flows each hour is less than 5%.

Table 5 Changes in passenger flows entering/exiting dwelling-oriented stations

Station type	Station name	Whole-day entering numbers	Whole-day exiting numbers	Entering rate during morning rush hours (%)	Exiting rate during morning rush hours (%)	Entering rate during evening rush hours (%)	Exiting rate during evening rush hours (%)
Dwelling-oriented	Nanmen	3682	3598	24	12	10	16
	Xiangmen	4933	4635	24	7	6	18



(a) Nanmen Station (b) Xiangmen Station

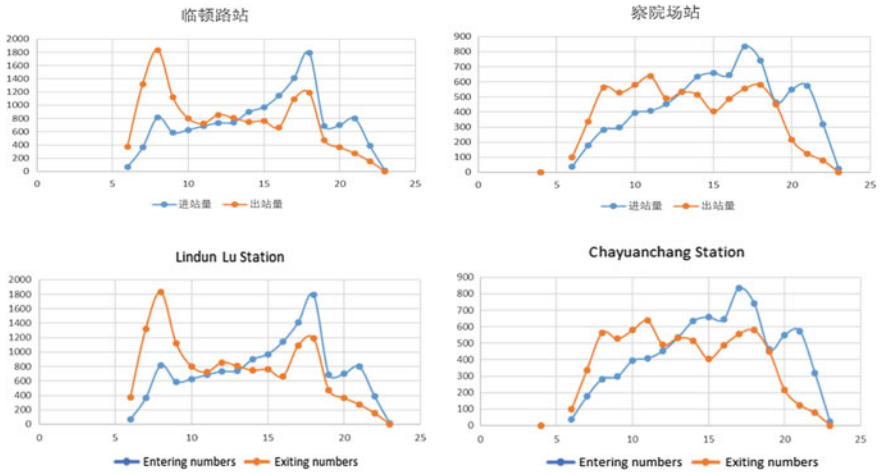
Fig. 2 Changes in passenger flows entering/exiting dwelling-oriented stations

(2) **Shopping-oriented Stations**

Hourly changes in passenger flows on weekdays at shopping-oriented stations inside the Area are analyzed. As is clear from Table 6 and Fig. 3, the "peak effect"—extremely high exiting rate during morning rush hours and entering rate during

Table 6 Changes in passenger flows entering/exiting shopping-oriented stations

Station type	Station name	Whole-day entering numbers	Whole-day exiting numbers	Entering rate during morning rush hours (%)	Exiting rate during morning rush hours (%)	Entering rate during evening rush hours (%)	Exiting rate during evening rush hours (%)
Shopping-oriented	Lindun Lu	13,433	13,544	6	14	13	9
	Chayuanchang	8025	7183	4	8	10	8



(a) Lindun Lu Station (b) Chayuanchang Station

Fig. 3 Changes in passenger flows entering/exiting shopping-oriented stations

evening rush hours—exists, albeit not obvious. In fact, both rates are less than 15%. And no huge differences emerge in the entering/exiting rate. During non-rush hours, the ratio of hourly entering numbers/exiting numbers stands at 6–8% or so. Worth noting is a small peak of entering passenger flows at 8:00 p.m.–9:00 p.m., which account for 6–7% of the total.

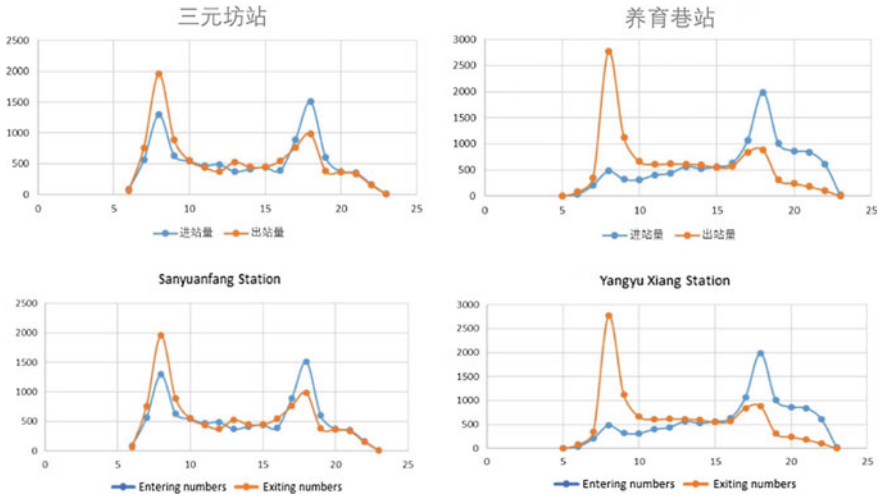
Shopping-oriented stations have three peaks, i.e. morning rush hours, evening rush hours and 8:00 p.m.–9:00 p.m. The third travel peak occurs as a large number of shoppers go home after commercial complexes and large smalls shut during that period.

(3) **Office-oriented Stations**

Table 7 and Fig. 4 describe the result of an analysis of hourly changes in passenger flows on weekdays at two office-oriented stations inside the Area. For office-oriented stations, during rush hours in the morning, most riders are exiting, and representing

Table 7 Changes in passenger flows entering/exiting office-oriented stations

Station type	Station name	Whole-day entering numbers	Whole-day exiting numbers	Entering rate during morning rush hours (%)	Exiting rate during morning rush hours (%)	Entering rate during evening rush hours (%)	Exiting rate during evening rush hours (%)
Office-oriented	Sanyuanfang	9634	10,011	13	20	16	10
	Yangyu Xiang	10,899	11,124	4	25	18	8



(a) Sanyuanfang Station (b) Yangyu Xiang Station

Fig. 4 Changes in passenger flows entering/exiting office-oriented stations

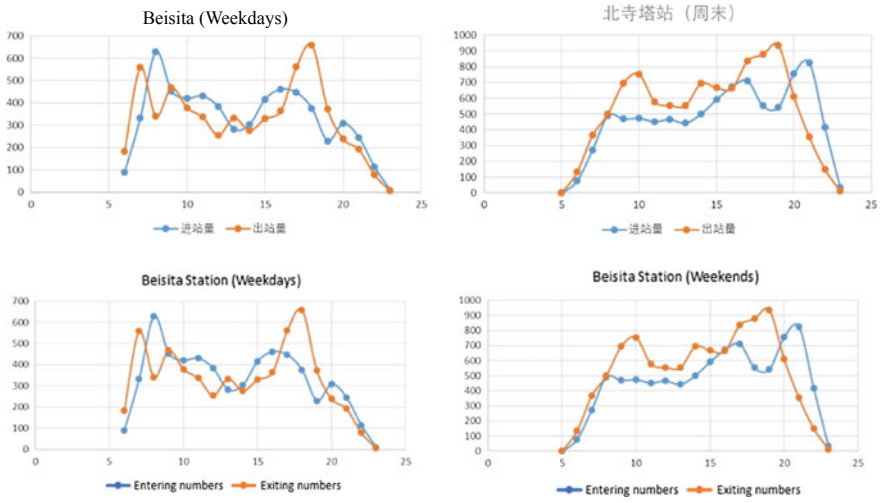
approximately 25% of the total, exiting ridership is more than four times higher than entering ridership; during rush hours in the evening, most riders are entering. Entering ridership, accounting for 15% of the total, is over twice that of exiting ridership. During the remaining non-rush hours, the proportion of hourly passenger flows entering/exiting the stations is below 5%.

(4) **Tourism-oriented Stations**

Characteristics of passenger flows at tourism-oriented stations vary depending on whether a rider travels on weekdays or on weekends. Therefore the paper takes a look at respective passenger flow changes separately, with the analysis results shown in Table 8 and Fig. 5. Beisita Station has a trend similar to the aforesaid typical shopping-oriented stations. On weekdays, entering numbers exceed exiting numbers during morning rush hours, and exiting numbers exceed entering numbers during evening rush hours. What happens on weekends is a little different. The station sees

Table 8 Changes in passenger flows entering/exiting tourism-oriented stations

Station type	Station name	Whole-day entering numbers	Whole-day exiting numbers	Entering rate during morning rush hours	Exiting rate during morning rush hours	Entering rate during evening rush hours	Exiting rate during evening rush hours
Tourism-oriented	Beisita	6355	5860	10%	5.7%	6.3%	11.1%



(a) Beisita Station (Weekdays) (b) Beisita Station (Weekends)

Fig. 5 Changes in passenger flows entering/exiting tourism-oriented stations

a concentration of passengers at the closing time of scenic spots, as evidenced by a surge in riders entering the station at 5:00 p.m. when the Beisita attraction closes.

5 Conclusion

The paper first explains the nature of different passenger flow collector-distributor points. Supported by rail transit AFC data, a case study is then conducted to give an insight into the characteristics of passenger flows at stations in Ancient Urban Area of Suzhou, which fall into dwelling-, shopping-, office- and tourism- oriented stations. The research results may apply to a range of fields concerning urban planning, including master planning, integrated transport planning, public transport planning, and rail transit planning.

References

1. Wang W, Zhang L (2019) Identification and classification method of rail transit station-influenced areas based on walkability index and network analysis—a case study of Tianjin. City 4:63–74 (in Chinese)
2. Wei Y, Tian Q, Guo T (2013) An improved pedestrian detection algorithm integrating haar-like features and HOG descriptors. Adv Mech Eng 1–8 (SCI)
3. Zhang X, Xu C (2018) A probe into layout and classification of urban rail transit central stations in Shenzhen. Urbanism Archit 17:51–54 (in Chinese)

4. Hong T, Jiang L (2018) Research on classification of rail transit stations in downtown Nanjing. Hous Real Estate 12:217 (in Chinese)
5. Chen Y, Zhang N, Lu Y (2017) A new classification approach to urban rail stations based on external connections. Urban Rapid Rail Transit 30(5):51–55 (in Chinese)
6. Duan D, Zhang F (2013) Study on classification of urban rail transit stations from the perspective of land use optimization: a case study on Xi'an Subway Line 2. City Plann Rev 37(9) (in Chinese)
7. Wei Y, Gao G (2015) Study on urban rail transit passenger flow crowd index. Urban Rapid Rail Transit J 28(2):5–9 (in Chinese)

Realized Application of a Contactless CPU Card for Public Transport Interconnection

Guo-jing Xing

Abstract Public transport cards from various cities cannot achieve transactions in other cities. This means they fail to meet current technological development trends and public travel demands. This paper puts forward a method to realize contactless CPU cards that feature cross-regional, cross-transportation-modes, as well as cross-industry interconnections in public transport. It does so by unifying the public transport card system composition, card files and application requirements, terminal transaction flows, clearing and settlement information interfaces, as well as the contactless interface communication, security system, test system. This method features compatibility between advanced financial electronic cash applications with traditional public transport card electronic wallet applications. It also features dual application shared files and balances, the national transport One-Card clearing and settlement platform building, as well as cards with international IIN and RID codes, the compatibility of a dual key system and dual algorithms, dual-coin function support. At present, more than 200 cities in China have joined the transport One-Card interconnection. More than 12 million One-Card interconnection cards have now been issued, and the national transport One-Card clearing and settlement platform has processed 30 million data exchanges after more than two years of application and implementation. The successful realization of the national transport One-Card interconnection has been improving the standards of public travel service and management in the transportation industry. It has also been useful for promoting the healthy development of the transport One-Card industry.

Keywords Contactless CPU card · Cross-regional and cross-transportation-mode interconnection · Dual application shared file and balance · International IIN code and RID code · Dual key system and dual algorithm compatibility

G. Xing (✉)
China Transport Telecommunications & Information Center, No. 1 Waiguan Houshen, Chaoyang District, Beijing, China
e-mail: xingguojing@cttic.cn

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_111

1 Introduction

In recent years, contactless CPU cards with high security, large storage space, fast reading speeds, and offline transaction capabilities have come to the forefront. Those supporting multiple applications with one card and multi-function development have become widely used within urban public transport One-Card in things like buses, the subway, taxis and other transportation modes. These application functions have also expanded to the field of small consumption payments [1–4]. However, there are different ways to realize the application of public transport cards among different cities. There are also a variety of terminals and card products in existing markets. International and domestic manufacturers are numerous, and the testing for terminals and cards is not strict. There are many problems with this, including the compatibility between terminals and cards, and the difference between transaction flows. Cross-regional and cross-transportation-mode interconnections cannot be achieved easily, and the current needs of big data and other technological trends cannot be met. Therefore, we should standardize and unify the composition of our public transport card systems, card files and application requirements, as well as our terminal transaction flows, clearing and settlement information interfaces, contactless interface communication, security systems and test systems. This must be done by formulating a unified standard (Technical specification on IC cards for urban public transportation (JT/T 978-2015) [5]). Under this precondition of unifying standards and in accordance with the technical specifications of the industry, ensuring unification of the technical specifications and the application implementation is the only way to fundamentally solve all of the problems.

Convenient and secure mobile payments are one of the key development directions for the payment industry in the future. As a payment means closest to people's livelihoods, the urban public transport One-Card application has become one of the main mobile payment fields. This is largely because of the large scale of users and high user viscosity. At present, mobile payment and financial IC cards have been used effectively by the urban public transport One-Card. Therefore, if we want to develop the urban public transport One-Card further, we should take into account the actual needs and technical level of the current urban public transport One-Card industry. We should also consider the requirements for transport One-Card cross-industry expansion to the field of small payment. In this regard, we need to formulate a standard that not only adopts the current electronic wallet application algorithms and flows, which are mostly used by local transport cards, but also one which is compatible with the algorithms and flows of financial electronic cash applications. Unifying and implementing compatible standards can provide technical support for the transport One-Card. This can help it to achieve cross-industry interconnection.

The required standard features the compatibility of advanced financial electronic cash applications [6, 7] with traditional public transport card electronic wallet applications [8], as well as dual application shared files and balances. It works toward building the national transport One-Card clearing and settlement platform, and establishing a test certification system, which provides sufficient technical guarantees that

allow the transport One-Card to realize cross-regional, cross-transportation-mode and cross-industry interconnection. The standard card features the adoption of international IIN and RID codes, and support for a dual currency function, which can support the transport One-Card interconnection cards to expand internationalization as well as lay a good foundation for transport One-Card interconnection to the international market. The technology characteristics of the standard feature the compatibility of a dual key system and the dual algorithms support the development of China's national cryptographic algorithm. A smooth transition from traditional technology to advanced technology is being now realized, and the safety and applicability of the cards is being enhanced.

With technical support from the standard, the construction of a transport One-Card interconnection has been carried out nationwide. The transport One-Card interconnection has now been realized in 181 cities at or above prefecture-level and 89 county-level cities. This means it covers Beijing-Tianjin-Hebei, the Yangtze River Delta, the Pearl River Delta and the Yangtze Economic Belt, as well as other areas following more than two years of promotion and implementation. More than 12 million transport One-Card interconnection cards have now been issued, and the national transport One-Card clearing and settlement platform has stably processed 30 million data exchanges. On the one hand, realization of the transport One-Card cross-regional, cross-transport-mode and cross-industry interconnection has greatly improved the standard of public travel services within the transportation industry. The public is generally very satisfied and believe that the transport One-Card interconnection improves the convenience of passengers travelling to different places, shortens ticket purchase time, and reduces transportation expenses. On the other hand, the realization of big data information analysis for national public travel has provided a new way to promote efficient technical management for management departments of the transportation industry and improved the overall industry management level. Meanwhile, enormous economic benefits have also been brought to production and operation enterprises, and healthy development of the transport One-Card industry has been promoted.

2 System Architecture

2.1 Urban Public Transport IC Card System Architecture

The urban public transport IC card system features systematic engineering. It is able to meet the different needs for public transport travel in different regions and the special business needs of different modes of transportation. It is compatible with payment standards for advanced financial electronic cash applications and traditional public transport card electronic wallet applications. Furthermore, it effectively supports standardized and automated management for transactions between different transportation modes, and ensures timely clearing and settlement of transaction data.

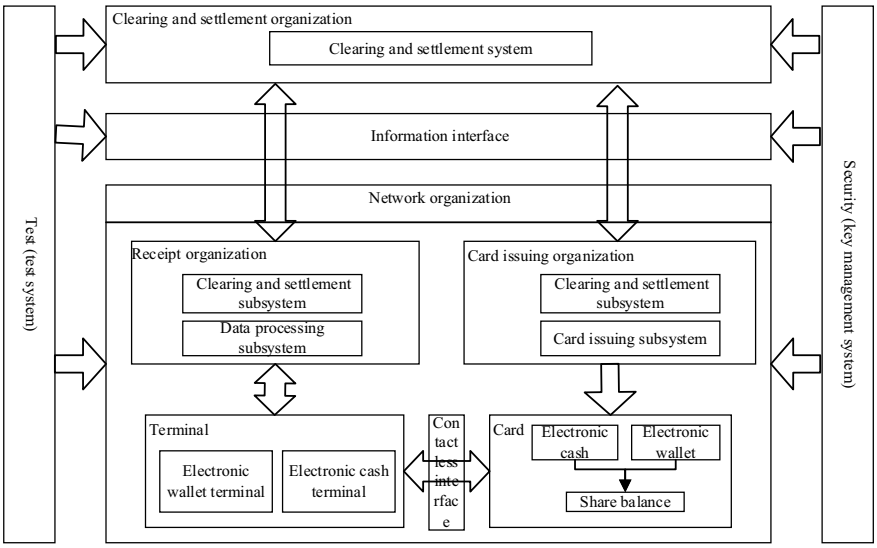


Fig. 1 Urban public transport IC card system architecture

The urban public transport IC card system is composed of a card, terminal, card issuing process, data processing, clearing and settlement, security, testing subsystems. The relationship between the subsystems is described below. First, consumption, loading, unloading and other functions are realized with the card as the carrier and terminal as the base. The timely circulation of transaction data between network organizations is done on a safe, reliable network with unified information interfaces. The clearing and settlement of the transaction data across organizations is completed by the clearing and settlement system according to clearing and settlement rules. The architecture of the urban public transport IC card system is shown in Fig. 1.

2.2 National Transport One-Card Interconnection System Architecture

In order to realize the interconnection of public transport cards within national network organizations, it is necessary to ensure normal transactions with cards at terminals in other cities. This must be done in accordance with standard transaction flows, and the clearing and settlement of the data must be uploaded to the unified data exchange platform following the transaction. Therefore, the national transport One-Card interconnection system is deployed with three levels of architecture, including the central system, the direct connection system and the indirect connected system. This is shown in Fig. 2. Among these, the national transport One-Card interconnection clearing and settlement platform ("The national transport One-Card platform"

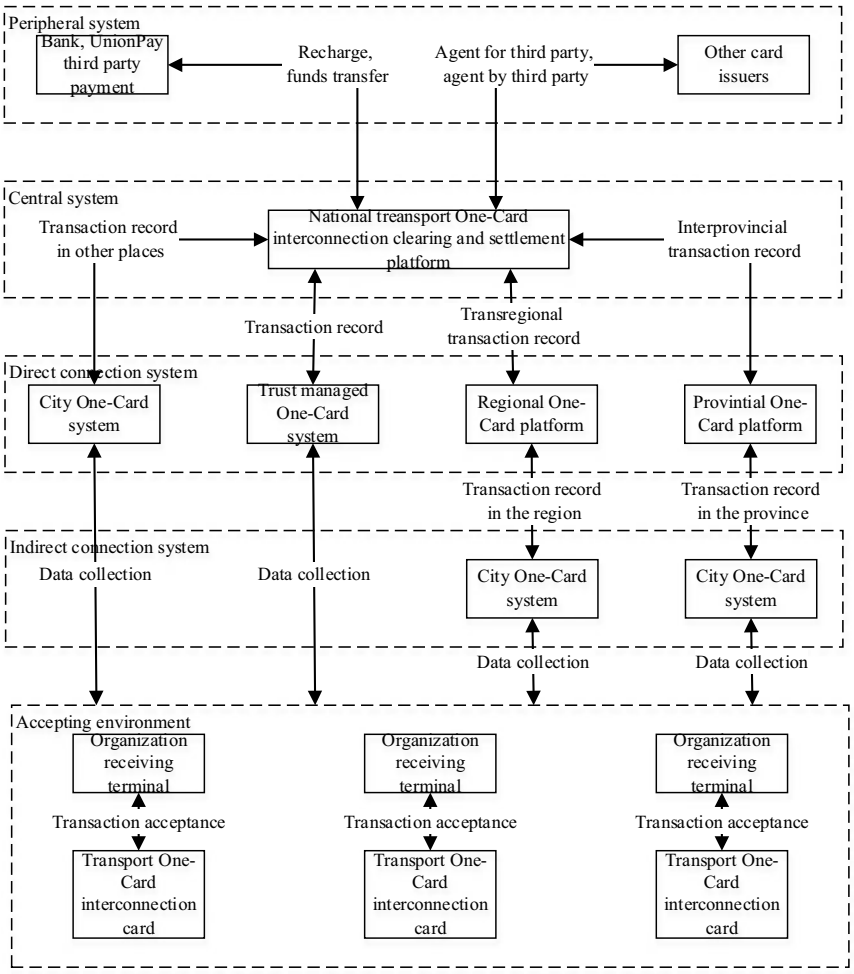


Fig. 2 National public transport card interconnection system architecture

for short) is responsible for the interconnection of the whole system. It is also responsible for managing key security over the whole system, clearing and settlement of national transaction data and any error adjustments. The provincial One-Card platform is responsible for clearing and settlement of transaction data within the province, as well as uploading the transaction data of cards issued by other province to the national transport One-Card platform for clearing and settlement in accordance with the unified information interface. The regional One-Card platform is responsible for clearing and settlement of transaction data within the region. It is organized according to economic development planning. Transaction data of cards issued by other region is processed in accordance with the unified information interface and uploaded to the

national transport One-Card platform for clearing and settlement. The city One-Card system handles all business from the public transport card system within the city, as well as uploads the transaction data of cards issued by other city in accordance with the unified information interface by directly and indirectly connected to the national transport One-Card platform for clearing and settlement. The trust managed One-Card system runs in the general centre, directly carrying out data switching, clearing and settlement.

3 Card Applications and File Requirements

In order to realize the interconnection of public transport cards for nationwide network organizations, we need to ensure that all the cards issued have unified application and file requirements. In order to make full use of the existing financial terminal environments and achieve cross-industry development, the transport One-Card interconnection cards adopt dual application technical specifications for electronic cash and electronic wallets. Two applications have their own independent areas, as well as shared areas. The card structure is shown in Fig. 3. In order to optimize the cardholder's experience, both applications share the same balance. At the same time, public transport process information record files are also shared between the two applications. Application data is recorded in time-by-time and section-by-section deductions in order to support the interconnection of electronic cash and electronic wallet terminals. This helps to achieve cross-regional and cross-transport-mode interconnection transactions.

The transport One-Card card features dual applications for electronic cash and electronic wallets. This determines that the card adopts both a symmetric key system and asymmetric key system. Meanwhile, the card should support switching between international and national cryptographic algorithms in order to support the development of a national cryptographic algorithm but still be able to consider existing international algorithms. Using this method, we can realize a smoother transition from traditional technology to more advanced technology, and enhance the security and applicability of the card at the same time.

In order to standardize the issuance of transport One-Card interconnection cards, we use the International issuer registration identification number (IIN section: 31047300-31052299) as the first eight digits of the card number. We also use the International application service provider registration identification number (RID: A000000632) as the two application's AID. These IIN and AID numbers are allocated by the International Organization for Standardization for use during the transport One-Card interconnection. This implementation supports the transport One-Card and allows it to be used internationally. In addition, in order to meet the needs of interconnections to Hong Kong, Macao and Taiwan in the future, the card electronic cash application supports a dual currency function. This function enables the card to support multi-currency electronic cash applications, and increases the overall applicability of the card.

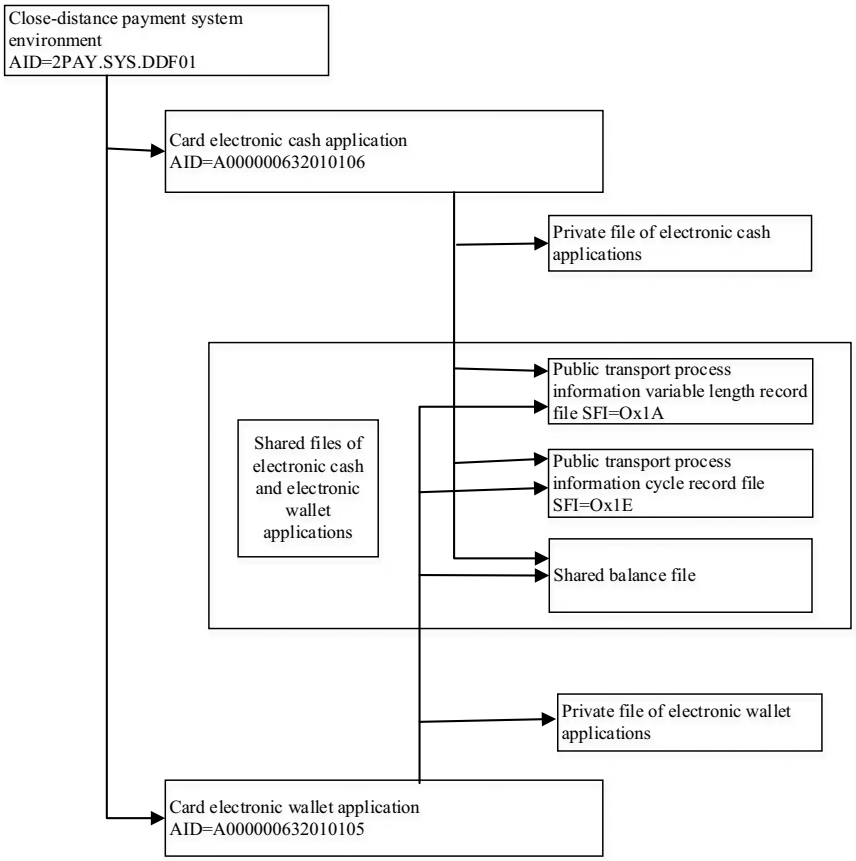


Fig. 3 Transport One-Card interconnection card structure

4 Terminal Transaction Flow

To realize transport One-Card interconnection, it is necessary to ensure that the terminals used by card issuers feature unified card discrimination and consumption transaction flows. In order to meet the industry characteristic of fast transportation, offline transactions from interconnection terminal consumption are used to support electronic cash applications and electronic wallet applications. According to different ways of deducting expenses, consumption transactions can be divided into standard fast payment transactions, time-by-time and section-by-section deducting transactions. They can also use off-line pre-authorized transactions, single discount transactions and other deduction transactions. The standard fast payment transaction is the most basic function. It is suitable for single deductions at fixed amounts (like those commonly used for public transport). The other payment functions are based

on the standard fast payment transaction. The time-by-time and section-by-section deducting transactions increase reading and updating for the public transport process information record file commands. This is based on the standard fast payment transaction flows, and is suitable for the secondary deduction scenes (such as with the subway). Here, the transaction amount is determined according to the time, place and other data found within the record files. The offline pre-authorized transaction is a special form of electronic cash, time-by-time and section-by-section deduction transaction. It is suitable for secondary deduction scenes (such as with a car rental) where there may be some type of deposit required. The single discount transaction is suitable for use with senior's cards, student cards and other discount scenes. Figure 4 illustrates the time-by-time and section-by-section deducting transaction flows for One-Card interconnection terminals with supporting electronic cash applications and electronic wallet applications, respectively. It uses a subway entry and exit as an example, as shown in Fig. 4.

Some incomplete transactions may occur during the time-by-time and section-by-section transactions and these make normal public transport card use in different places impossible. In order to solve these problems, we put forward a supplementary deduction scheme for the incomplete transactions that may be used by the One-Card interconnection card in different places. By taking the subway scenarios for an example, the transport One-Card interconnection terminal determines whether the last transaction in the card is incomplete during entry by reading the public transport process information variable length record file (before exit). If the transaction is incomplete, the terminal can automatically deduct a certain amount according to the information it discovers. It can also update the transaction status to exit, and then the card can start the next normal transaction. The transaction record produced by deduction in other places will then be uploaded to the national transport One-Card platform. This means that clearing and settlement with the card issuers, the receiving organizations, and the local organizations occur even with incomplete transactions in accordance with certain rules.

5 National Transport One-Card Interconnection Clearing and Settlement Platform

To achieve a nationwide interconnection of the transport One-Card, and in addition to the normal transaction of the card in terminals found in other places, the national transport One-Card interconnection clearing and settlement platform needs to be built. This is needed for storing and forwarding transaction data uploaded by all the national network organizations. It is also needed for daily batch-completing clearing and settlement, distributing the results of clearing and settlement, adjusting the transaction errors, and for analysing national transport big data. The national transport One-Card platform is also responsible for managing key security for the entire system. According to the relationship between the system data flow and business flow, the

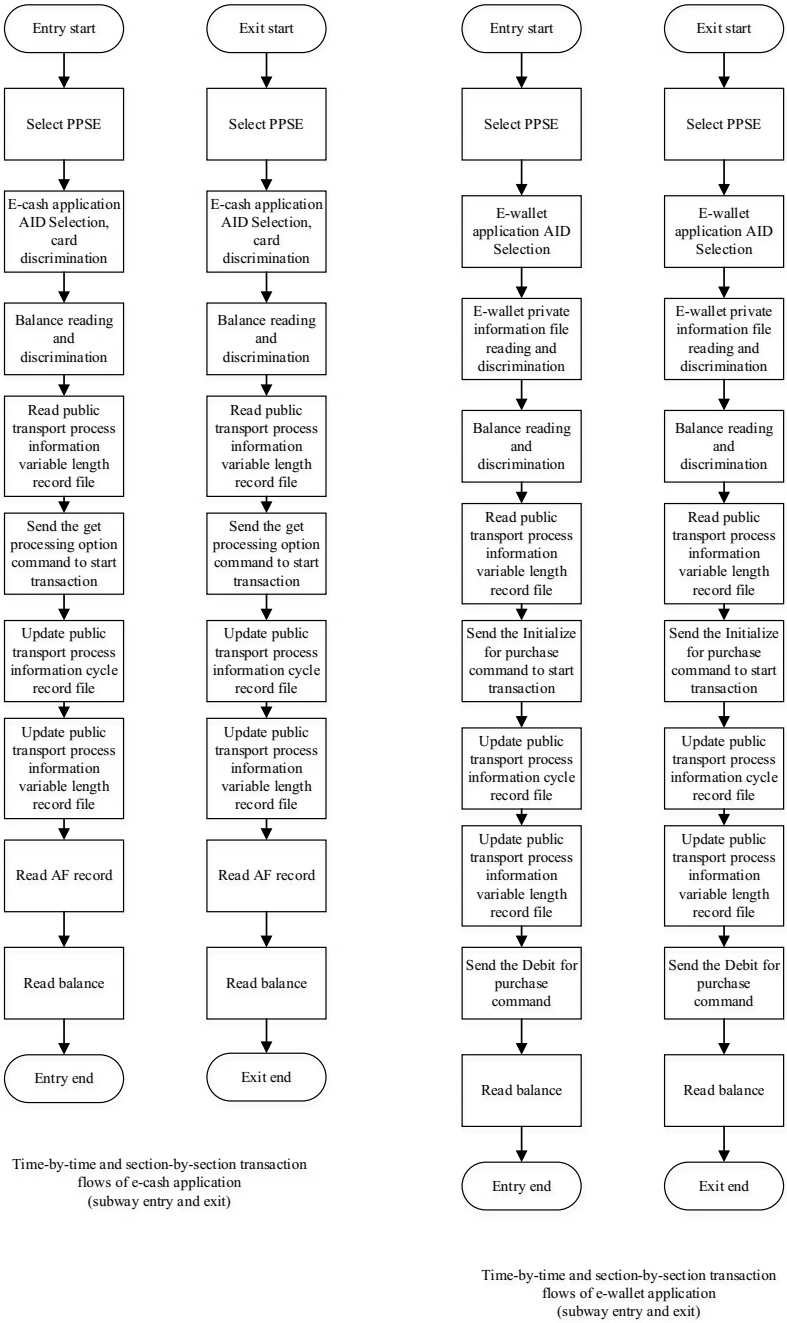


Fig. 4 Time-by-time and section-by-section deducting transaction flows for One-Card interconnection terminals

clearing and settlement interfaces can be divided into: communication interfaces, transaction data interfaces, accounting data interfaces, parameter data interfaces, credit data interfaces, error handling interfaces, interfaces with the clearing banks, as well as interfaces with network organizations.

According to the concept of open development, the national transport One-Card interconnection system can accept cards issued by other card issuers. For the clearing and settlements during interconnection with other card organizations, the information interface specification from this system can be used for file transfer and accounting. It also supports the information interface specifications defined by the other card organizations for file interactions.

6 Test Certification System

To ensure the smooth realization and sustainable development of national transport One-Card interconnection, it is necessary to develop a set of complete and detailed test specifications. These must also be compatible with the technical specifications. On this basis, steps should be carried out to develop test tools, establish the test certification system, and to strictly manage cards, terminal products and systems used in the transport One-Card interconnection. This allows us to fundamentally solve the terminal-card compatibility, transaction flow differences and other common problems related to transport One-Card interconnection. It also allows us to realize cross-regional and cross-transportation-mode travel by means of just one transport card. In this way, greater convenience and flexibility will be provided for public travel.

The test specifications determine the methods and requirements of the card test, SAM card test and terminal test, respectively. They do so according to three levels: physical characteristics, contactless electrical characteristics and communication protocols, as well as any upper applications. The test requirements needed to correctly realize information interfaces at the system level are put forward here. For these test specifications, corresponding test cases are innovatively increased on the basis of older test cases from electronic cash and electronic wallets found within the financial industry. They utilize dual application shared files and balance, dual application linkage locking, new parameter support for electronic wallet balance commands, electronic cash special data elements, as well as other specific functions from the One-Card interconnection specifications. This helps to ensure effectiveness and conformance for all the test specifications.

At present, the cards, terminal products and systems involved in the national transport One-Card interconnection have been developed and used in accordance with these test specifications. By taking the actual operations of the national transport One-Card interconnection system for more than two years, the rationality, effectiveness and feasibility of the test certification system has been verified, and the application requirements for interconnection have now been satisfied.

7 Innovation Points

The dual application shared file and balance technical route has been adopted by the transport One-Card interconnection cards. The application of the electronic wallet is based on the technical status of the transport One-Card industry. It saves on the costs of terminal transformation and protects the interests of the industry. The application of electronic cash is used to take into account the development direction of more advanced technology and realizes transport One-Card's cross industry development. Two applications share files and balances in order to enhance a user's experience. This means there is technological innovation, but also the advantages associated with achieving practical promotion.

The first eight digits of the card number use the international IIN code, and the two application AID numbers use the international RID code. This enables the card to be used internationally.

The card adopts a symmetric key system and an asymmetric key system at the same time, and it supports international and national double algorithm switching. This implementation realizes a smooth transition from traditional technology to more advanced technology, and also enhances the safety and applicability of the card. The corresponding key management system has now obtained a national product certificate issued by the SCA, which is the top national product certification in the transportation industry.

In addition to the standardization of the card and the terminal, we also innovatively built the national transport One-Card interconnection clearing and settlement platform. It can help to realize clearing and settlement of transaction data across the whole country.

In order to ensure that transport One-Card interconnection products strictly meet all technical specification requirements, we independently developed testing tools, established a testing certification system, and obtained the approval from the Certification and Accreditation Administration of the People's Republic of China (CNCA). This is the top inspection and certification authorization for transport card products.

We also registered the trademark CHINA T-UNION, which is used to identify all national transport One-Card interconnection products.

8 Concluding Remarks

This paper described a method for achieving contactless CPU card cross-regional and cross-transportation-mode interconnections for public transport. It did so by unifying standards for the transport One-Card card, terminal products and systems. This method has now been implemented for more than two years. This verifies that it can effectively solve terminal-card compatibility, differences in transaction flow and other common problems with the transport One-Card. More than 200 cities in China

have now realized the transaction of public transport cards. Through the implementation of this method, the realization of a national transport One-Card interconnection embodies significant economic and social benefits. In terms of economic benefits, the waste of resources associated with too many cards used is avoided. Furthermore, product competitiveness for the R&D and production enterprises is improved, and production and maintenance costs are also reduced. The large-scale use of the transport One-Card card also brings increasing economic benefits to operation enterprises. In terms of social benefits, the service standards of public travel have been improved by this method. Big data collected on national public travel also provides comprehensive data analyses and decision-making information for transportation industry management departments. In this way, the idea helps with improving the management standards of the transportation industry. Meanwhile, compatibility with advanced financial electronic cash applications, the adoption of international IIN and RID codes, as well as the dual currency functions and other technical characteristics of these cards lay a good foundation for achieving One-Card interconnection cross-industry applications and heading towards the international markets.

References

1. Standardization Administration of the People's Republic of China (2006) GB/T 14916-2006, Identification cards—physical characteristics. Standards Press of China, Beijing
2. Standardization Administration of the People's Republic of China (2010) GB/T 16649-2010, Identification cards—integrated circuit cards with contacts. Standards Press of China, Beijing
3. Standardization Administration of the People's Republic of China (2008) GB/T 17552-2008, Identification cards—financial transaction cards. Standards Press of China, Beijing
4. Standardization Administration of the People's Republic of China (2006) GB/T 17554-2006, Identification cards—test methods. Standards Press of China, Beijing
5. Ministry of Transport of the People's Republic of China (2015) JT/T 978-2015, Technical specification on IC cards for urban public transport. China Communications Press, Beijing
6. The people's bank of china (2005) JR/T 0025-2005, China financial integrated circuit card specifications. China Financial Publishing House, Beijing
7. The People's Bank of China (2014) JR/T 0045-2014, Specifications for detecting Chinese financial integrated circuits cards. China Financial Publishing House, Beijing
8. Ministry of Housing and Urban-Rural Development of the People's Republic of China (2007) CJ/T 243-2007, Test methods for construction cause IC cards. China Architecture & Building Press, Beijing

Bearing Fault Diagnosis with Impulsive Noise Based on EMD and Cyclic Correntropy

Yu-Ze Wang, Yong Qin, Xue-Jun Zhao, Shun-Jie Zhang and Xiao-Qing Cheng

Abstract Periodic pulses are an important fault feature of rolling bearings, so the ability to accurately and efficiently identify pulse components is important for bearing fault diagnosis. Due to the complicated wheel-rail contact relationship in actual train operation, it often generates many impulse noises which similar to the fault signal structure. Unfortunately, spectral kurtosis (SK) methods often fail to effectively diagnose under impulse noise. In order to solve this problem, this paper proposes a bearing fault diagnosis method based on Empirical Mode Decomposition (EMD) and cyclic correntropy (CCE) function. Compared with the SK method, the method proposed in this paper can effectively suppress the influence of impulse noise. Moreover, this paper also proposes a fault diagnosis evaluation index KR_s to quantitatively compare the diagnostic effects of different methods. Simulations and real data of the train axle are utilized to demonstrate the feasibility and effectiveness of the proposed method and index.

Keywords Train bearing fault diagnosis · Impulsive noise · Correntropy · EMD · Cyclostationary · Fault evaluation index

1 Introduction

Rolling bearings are the most widely used mechanical components for all types of rotating machinery. According to statistics, 30% of rotating machinery faults are caused by bearing faults [1]. The rolling bearing in the bogie of the railway train is also one of the wearing parts, and its working state directly affects the operational safety of

Y.-Z. Wang · Y. Qin (✉) · X.-J. Zhao · S.-J. Zhang · X.-Q. Cheng
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
e-mail: yqin@bjtu.edu.cn

Y.-Z. Wang
e-mail: 17120894@bjtu.edu.cn

Y.-Z. Wang · X.-J. Zhao · S.-J. Zhang
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_112

the train. Therefore, it is of great significance for the safety of railway transportation to accurately detect and identify bearing faults and analyze the causes, especially to explore diagnostic methods suitable for railway application background.

When a local defect occurs in the rolling bearing, the bearing will generate periodic pulses, so frequency domain analysis is the preferred method for diagnosing local defects of the bearing [2]. However, this useful fault signature information is often overwhelmed by strong background noise. Therefore, the key to fault diagnosis of rolling bearings is to extract the fault characteristic frequency from the modulated vibration signal. This proved to be an effective method that typically uses bandpass filtering combined with envelope techniques to filter out unrelated components and then extract and demodulate the signal [3, 4]. For example, spectral kurtosis (SK) [5, 6] and Kurtogram algorithms have proven to be very effective for quantifying pulse components in different frequency bands in a signal. The basic idea of SK is to use the kurtosis index to select the frequency band that is most likely to have a pulse component, and then perform a square envelope analysis after bandpass filtering. Under the influence of impulse noise, SK may cause erroneous selection of the optimal frequency band and bandwidth, which limits the performance of kurtogram in identifying mechanical faults [7]. With the new concept named Infogram [8], it has been proved that the superiority and the importance of cyclostationary analysis for suppressing impulse noise.

In 1986, Gardner proposed the concepts of cyclic autocorrelation and spectral correlation for studying cyclostationary signals, and established a spectral correlation theory system in subsequent research [9–11]. It provides a theoretical basis for the cyclostationary phenomenon that is widespread in various research fields. In the field of mechanical vibration analysis, the physical parameters of the rotating machine have the characteristics of periodic time-varying during the operation of the rotating machine, especially when there is a fault, showing cyclostationarity. Antoni [12] summarized the mechanism research of a variety of rotating machinery cyclostationary, systematically discussed the cyclostationary phenomenon and the generation mechanism of cyclostationary in rotating machinery. In 2017, a fast algorithm for spectral correlation analysis [13] was proposed, which greatly facilitated the study of bearing fault diagnosis based on spectral correlation analysis.

In recent years, the basic definition of stochastic process correlation functions has been extended to generalized correlation functions in the field of communication signal processing, which is called correntropy function. The correntropy function can handle non-Gaussian noise and impulsive noise well [14, 15]. Then, a cyclostationary analysis technique with correntropy function called CCE emerges [16]. The research shows that CCE has a good effect on carrier frequency estimation and time delay estimation under impulsive noise environment. The results show the suppression effect of correntropy on impulse noise, and introduce it into the fault diagnosis of rolling bearings to solve the diagnostic error under impulse noise conditions. However, in addition to impulse noise in the vibration signal, there are other components that we do not expect to see. Therefore, in this paper, the preprocessing is combined with the CCE. The EMD algorithm is used to filter

out the clutter, and the display of the fault information in the CCE analysis is enhanced to achieve the purpose of accurate fault diagnosis.

The rest of the paper is summarized as follows: In the second section, the basic principles of EMD algorithm, correntropy function and cyclostationary analysis are briefly introduced. In the third section, a cyclostationary analysis method based on correntropy entropy function and EMD and its spectrum for bearing fault diagnosis is proposed. In the fourth section, the effectiveness of the proposed method is verified by the simulated signal containing impulse noise. In addition, an evaluation index for the effectiveness of the fault diagnosis method is proposed. In the fifth section, two real fault bearing signals (including the real train fault bearing signal), are used to verify the effectiveness of the proposed method. And compared with the SK algorithm to highlight its superiority.

2 Fundamentals of Correntropy Function and Cyclic Spectral Analysis

2.1 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) [17] is an adaptive signal decomposition method, which decomposes the signal into multiple Intrinsic Mode Functions (IMFs) and one residual component $r(t)$ at the end by decomposing the signal into different characteristics of time scale or variation trend, as shown in Eq. 1. EMD, especially for non-stationary and nonlinear signals, can achieve a good decomposition effect. The decomposition process of EMD is to continuously extract the scale components of the original signal from high frequency to low frequency. The result of EMD decomposition is first a highest frequency component, then a sub-high frequency component, and finally a residual component with a frequency approaching 0.The high frequency components with high energy in these components often represent the main characteristics of the original signal. Therefore, EMD, as a principal component analysis method, can extract the main components of signals in priority.

$$x(t) = \sum_{i} imf_i(t) + r(t) \tag{1}$$

The IMF resulting from EMD decomposition need to satisfy the following two necessary condition:

(1) The extremum number of a single component should be equal to or at most one different from the zero number in the entire time domain.
(2) For a single component, it is locally symmetric on the time axis, which means the mean value of the upper envelope of the local maximum and the lower envelope of the minimum should be zero or close to zero.

According to the above IMF restrictions, the EMD algorithm can be described in detail as following steps:

(1) By obtaining the upper and lower extremum points of the original signal, the upper envelope $e_{upper}(t)$ and lower envelope $e_{lower}(t)$ were fitted by cubic spline interpolation.
(2) The mean envelope e_{mean} is calculated and subtracted from the original signal $y(t)$ to get a new data sequence $y_1(t)$, which low frequency removed.

$$e_{mean} = \frac{1}{2}\left(e_{upper}(t) + e_{lower}(t)\right) \tag{2}$$

$$x_1(t) = x(t) - e_{mean} \tag{3}$$

(3) If the new signal $x_1(t)$ satisfies the necessary conditions of IMF, then $imf_1(t) = x_1(t)$; if not, repeat steps 1 and 2 for $x_1(t)$ until the generated new signal satisfies with this IMF's necessary conditions, denoted it as $imf_1(t)$.
(4) Subtract $imf_1(t)$ from the original signal $y(t)$ to get a data sequence $r_1(t)$ with high frequency components removed, and repeat steps 1, 2 and 3 to get $imf_2(t)$. This is repeated until the last difference sequence $r_n(t)$ can no longer be decomposed. Thus, all IMF can be obtained. The result is shown as Eq. 1.

In an ideal state, the IMF obtained by EMD decomposition is a complete, adaptive and nearly orthogonal expression of the original signal. There should be a good correlation between the true IMF component signals and the original ones. However, the correlation between some of the false IMF signals caused by endpoint effects [18] and the original signals will be poor. In this paper, the correlation coefficient between IMF and original signal is taken as an evaluation standard to select effective IMF components [19].

3 Correntropy Function

Correntropy, as a measure of local similarity between random variables, has attracted much attention in recent years [15, 20–25]. Correntropy was initially defined as a random process [15], and later this concept was further extended. For two random variables, the correlation entropy is defined as,

$$C_\sigma(X, Y) = E[k_\sigma(X - Y)] \tag{4}$$

In this formula, $E(\cdot)$ is the mathematical expectation, $k_\sigma(\cdot)$ is the kernel function, and σ is the kernel width. The kernel function can map the input data from the low-dimensional space to the high-dimensional space, and the higher the number of dimensions mapped to the high-dimensional space, the better the performance of distinguishing the differences of the input data. Since the Gaussian kernel function

[26] can map the input data into an infinite dimensional space, so it has been selected in this paper. The expression of the Gaussian kernel function can be shown as follows,

$$k_\sigma(X-Y)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\|X-Y\|^2}{2\sigma^2}\right) \tag{5}$$

Substituting Eq. 5 into Eq. 4, the expression of correntropy based on Gaussian kernel function can be obtained as follow,

$$C_\sigma(X,Y)=\frac{1}{\sqrt{2\pi}\sigma}E\left[\exp\left(-\frac{\|X-Y\|^2}{2\sigma^2}\right)\right] \tag{6}$$

As for core width σ, the classic Silverman rule [26] is selected in this paper, as shown in Formula (7), which can reduce the calculation amount on the premise of guaranteeing the effect,

$$\sigma=0.9AN^{-1/5} \tag{7}$$

In the above formula, N is the data length, $A=\min\{\hat{\sigma}, R/1.34\}$, $\hat{\sigma}$ is the sample standard deviation, and R is the quartile moment of the sample data.

Correntropy is essentially a second order statistic of mapping feature space data. The correntropy function based on Gaussian kernel function has the following two advantages:

(1) As a measure of local similarity, correntropy can reduce the influence of abnormal points.
(2) Correntropy contains more information than traditional correlation functions. The following equation can be obtained by Taylor expansion of Gaussian kernel function,

$$C_\sigma(X,Y)=\frac{1}{\sqrt{2\pi}\sigma}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n\sigma^{2n}n!}E\left[\|X-Y\|^{2n}\right] \tag{8}$$

It can be seen that the expression of correlation entropy is the sum of all even moments of two variables X and Y. Therefore, the correlation entropy function not only contains the information of cross-correlation function, but also has more information than the traditional correlation function.

4 Cyclic Spectral Analysis

Cyclostationary signal is a special kind of non-stationary signal, whose statistical characteristics are periodic [9, 12]. Suppose signal $x(t)$, $y(t)$ are all cyclostationary signals with the same periodicity, then the instantaneous cross-correlation function of signals is as follow,

$$R_{xy}(t,\tau) = E\left\{x(t)y^*(t+\tau)\right\} = R_{xy}(t+T,\tau) \tag{9}$$

where, $E(\cdot)$ is the mathematical expectation, * means the complex conjugate, and τ means the delay time,

Because $R_{xy}(t,\tau)$ is a function of period T, $R_{xy}(t,\tau)$ can be expressed in the form of Fourier series,

$$R_{xy}(t,\tau) = \sum_{\alpha} R_{xy}^{\alpha}(\tau)e^{j2\pi\alpha t} \tag{10}$$

where, α refers to Cyclic frequency; $\alpha = m/T$ and $R_{xy}^{\alpha}(\tau)$ to Cyclic cross-correlation Function (CAF); Fourier coefficient $R_{xy}^{\alpha}(\tau)$ indicates the amplitude of instantaneous autocorrelation $R_{xy}(t,\tau)$ in frequency α. By performing Fourier transform to the time delay τ of Cyclic Cross-correlation Function $R_{xy}^{\alpha}(\tau)$, you can get Cyclic Spectrum Density Function (CSD) as follow,

$$S_{xy}^{\alpha}(f) = \frac{1}{T}\int_{-\infty}^{\infty} R_{xy}^{\alpha}(\tau)e^{-j2\pi f\tau}d\tau \tag{11}$$

$S_{xy}^{\alpha}(f)$ is also known as spectral correlation, or spectral correlation function, which represents the circular spectral density value of a certain frequency f in the spectrum of circulatory stationary signal, and is a dual-frequency plane function about frequency f and circulatory frequency α. When a = 0, $R_{xy}^{0}(\tau)$ is the signal's cross-correlation function $R_{xy}(\tau)$, and $S_{xy}^{0}(f)$ is the signal's cross-power spectral density $S_{xy}(f)$.

Instantaneous cross-correlation hides the signal from periodic energy accumulation. Spectral correlation takes cyclic cross correlation as the bridge, and after two Fourier transforms, the two-dimensional periodic instantaneous cross correlation gradually converges into the peak, the spectral peak in the spectral correlation graph is the periodic performance of the signal, and the attribute of spectral peak reflects the characteristic of periodicity.

5 Cyclostationary Analysis Based on EMD and Cyclic Correlation Entropy

Assume that the correntropy function of $x(t)$ and $y(t)$ of random processes with the same period T_0 is $C_{xy}(t,\tau)$. According to the calculation process of spectral correlation function, such as Formulas (9)–(11), the cyclic correntropy function $C_{xy}^{\alpha}(\tau)$ and cyclic spectral density function $S_{xy}^{\alpha}(f)$ based on the correntropy function can be obtained,

$$C_{xy}^{\alpha}(\tau) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} C_{xy}(t,\tau)e^{-j2\pi\alpha t}dt \tag{12}$$

Combined with Formula (4), the cyclic correntropy function $C_{xy}^{\alpha}(\tau)$ can be denoted as

$$C_{xy}^{\alpha}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} k_{\sigma}(x(t), y(t+\tau)) e^{-j2\pi\alpha t} dt \tag{13}$$

$$S_{xy}^{\alpha}(f) = \int_{-\infty}^{\infty} C_{xy}^{\alpha}(\tau) e^{-j2\pi f\tau} d\tau \tag{14}$$

$S_{xy}^{\alpha}(f)$ shows the distribution of frequency spectrum f and cycle frequency α. The carrier refers to the spectral band of frequencies and the modulation refers to the cyclic frequency. The cycle frequency α characterizes the modulation, i.e. the fault frequency of the bearing can be seen.

Based on the above introduction, the specific steps of cyclic stationary analysis based on EMD and correntropy function are shown as follows,

Step 1. Obtain bearing vibration data $x[n]$, where n is the signal length;
Step 2. Perform EMD decomposition of the original signal $x[n]$ to obtain several IMF, and select the IMF with the highest correlation with the original signal as $y[n]$;
Step 3. Calculate the correntropy function $C_{xy}[\tau_n]$ of signals $x[n]$ and $y[n]$, and on this basis calculate the Fourier transform result $C_{xy}^{\alpha}[\tau_n]$ of the correntropy function in α domain;
Step 4. Calculate the Fourier transform result $S_{xy}^{\alpha}[f]$ of cyclic correntropy $C_{xy}^{\alpha}[\tau_n]$ in f domain;
Step 5. In order to identify faults more clearly, $S_{xy}^{\alpha}[f]$ is projected in domain α to obtain the spectral distribution results of fault signals.

6 Simulation Analysis

In general, a series of pulse vibration signals will be generated when local faults occur in rotating mechanical parts in operation. According to reference [27], the simulated bearing fault signals modeled by a single-degree-of-freedom system can be expressed as

$$y(t) = \sum_{k} \exp[-\beta(t - kT - \tau_k)] \times \sin[2\pi f_{re}(t - kT - \tau_k)] \tag{15}$$

where, β is the structural damping ratio, T is the pulse interval, and its inverse is the fault frequency f_c, f_{re} is the resonance frequency, and τ_k obeys the uniform discrete distribution to simulate the random slip caused by roller sliding.

Impulse noise is a kind of impulse signal with similar structure to bearing fault signal, which can also be modeled as a single-degree-of-freedom system, but with different structural parameters. The parameter Settings of bearing fault signal $y_1(t)$ and impulse noise signal $y_2(t)$ are shown in Table 1.

Table 1 Parameters of the simulated signal

	β	T	f_c	f_{re}
$y_1(t)$	900	0.01	100	1000
$y_2(t)$	300	1/30	30	3000

In this section, in order to verify the identification ability of different methods for bearing fault signals containing impulsive noise, as shown in Formula (16), the simulated signal $h(t)$ is mainly composed of three parts, the bearing fault signal $y_1(t)$ and the impulsive noise signal $y_2(t)$ and random white noise $n(t)$.

$$h(t) = y_1(t) + y_2(t) + n(t) \tag{16}$$

The sampling frequency F_s and sampling point N are set as 10,000 Hz and 10,000 respectively, and the signal-to-noise ratio of $h(t)$ is -6 db. Impulse noise starts from the 1000th sampling point and ends with 9 consecutive impulses. The amplitude of impulse noise is significantly higher than the bearing fault signal. The time domain diagram of simulated signal is shown in Fig. 1, where Fig. 1a is the fault signal with impulse noise, and Fig. 1b is the final simulated signal $h(t)$.

The method of SK [5] is to create a bandwidth-center frequency array, conduct bandpass filtering on time-domain signals and calculate the envelope, and select the optimal filtering bandwidth and center frequency by calculating the kurtosis index of each time-domain envelope signal. The corresponding Kurtogram is obtained by Fast Kurtogram [6]. As shown in Fig. 2, it is obvious that the region with the maximum kurtosis is around 3000 Hz, which is the resonance frequency of impulse noise nearby, while the fault features with small amplitude are hidden by impulse noise with large amplitude. Through band-pass filter to get the center frequency $f_c = 2968\,\text{Hz}$,

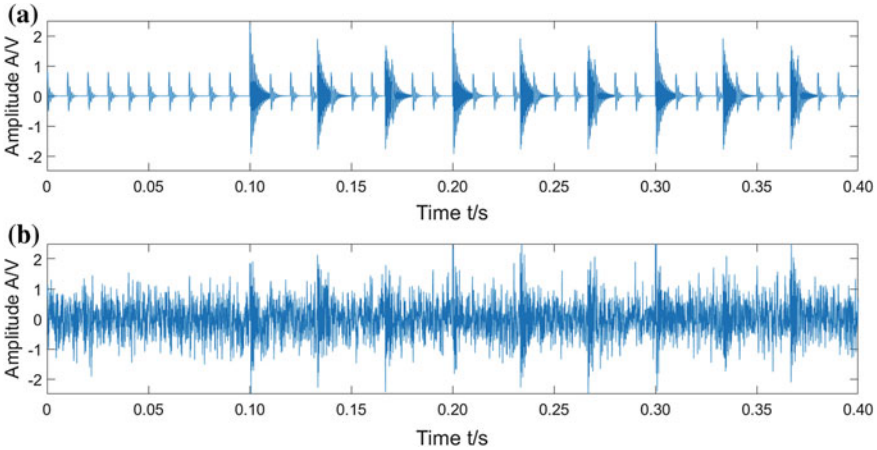


Fig. 1 **a** $y_1(t) + y_2(t)$; **b** 模拟信号 $h(t)$ (SNR $= -6$ dB)

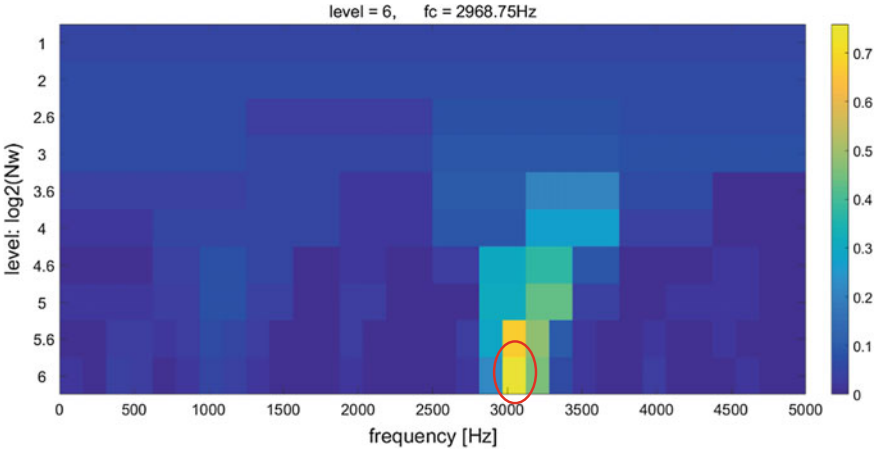


Fig. 2 Fast Kurtogram of the simulated signal

bandwidth $B_w = 156\,\text{Hz}$ filtering signal, as shown in Fig. 3a, its corresponding Hilbert spectral envelope of square as shown in Fig. 3b. From Fig. 3b, we can see that the optimal frequency band selected by the method SK is actually the frequency band where the impulse noise component is located. The fault characteristic frequency in the figure is 30 Hz, which corresponds to the characteristic frequency of impulse noise, the bearing fault frequency we expect to see is completely masked.

Furthermore, the algorithm proposed in the third section is used to analyze the characteristics of the simulated signal, and the projection of the simulated signal on

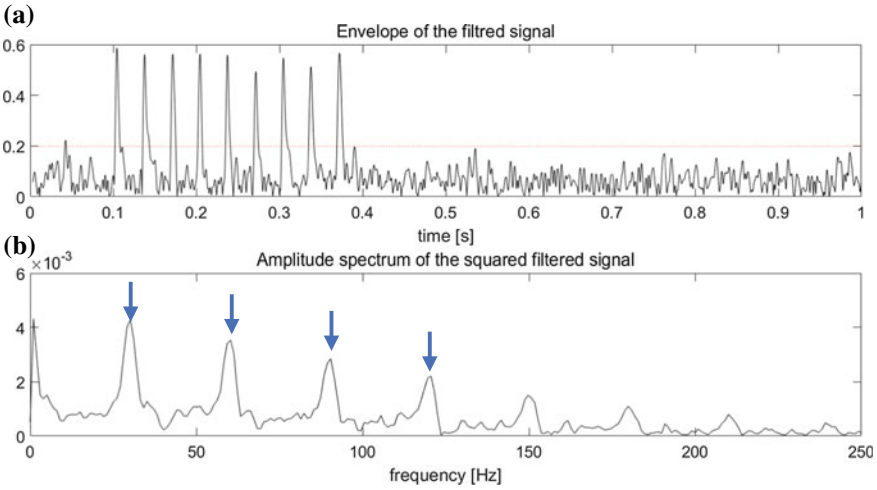


Fig. 3 **a** Envelope of the filtered signal which maximizes the Kurtogram. **b** Amplitude spectrum of the squared envelope

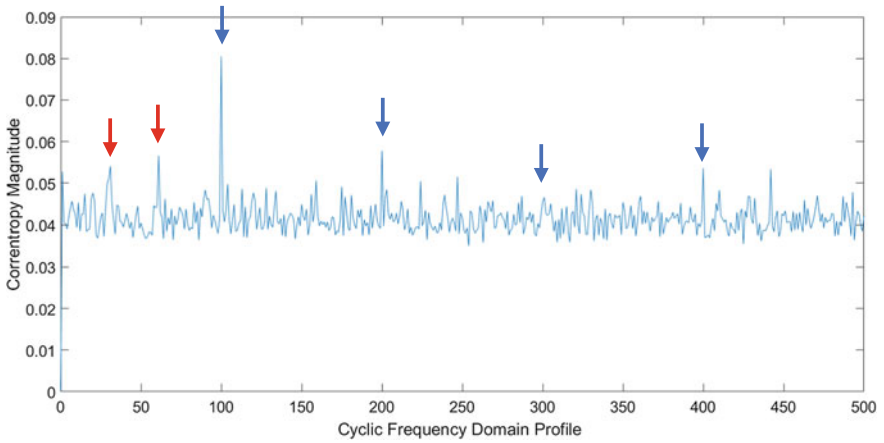


Fig. 4 Cyclic domain profile of the simulated signal

the cyclic frequency domain is obtained. The specific results are shown in Fig. 4. As you can see from the figure, the desired 100 Hz and its harmonics (fault frequency) are easy to identify, marked with blue arrows in this picture. At the same time, the spectrum in the figure also shows the frequency component of 30 Hz (impulse noise). Compared with the fault frequency of 100 Hz and its harmonics, impulse noise has been greatly suppressed. Therefore, the analysis results of the simulated signal under the influence of impulse noise show that the proposed method can effectively detect the fault frequency.

In order to more clearly and accurately compare the performance of these methods, a dimensionless index KR_s is introduced, which consists of the product of a dimensionless index K (Kurtosis) [1] and envelope spectrum fault feature ratio R_s,

$$KR_s = K \times R_s \tag{17}$$

For signal $x(t)$, the two parameters are calculated as,

$$K = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \overline{X}}{\sigma_x}\right)^4 \tag{18}$$

$$R_s = \frac{\sum S(f)}{S} \tag{19}$$

where, $\overline{X}$ and σ_x are the mean value and standard deviation of x, N is the signal length, f is the characteristic frequency of bearing fault, $S(f)$ is the amplitude of envelope spectrum at the characteristic frequency, and S is the total amplitude of envelope spectrum.

Table 2 The fault feature indexes of the simulated signal using different methods

Simulation signal	Proposed method	SK
K	27.677	9.657
R_s	0.011	0.004
KR_s	0.311	0.041

Specifically, index K can be used to represent the number of interference frequencies in the envelope spectrum. The larger the value of K, the less interference except the fault characteristic frequency in the envelope spectrum, and vice versa. Index R_s represents the prominence of fault characteristic frequency in the square envelope spectrum. The larger R_s value is, the more obvious the fault characteristic frequency. Then, the index KR_s is used to evaluate the performance of different fault diagnosis methods. The larger the value of exponential KR_s is, the more obvious the fault characteristic frequency is in the square envelope spectrum, the less the interference frequency is, and the better the algorithm performance.

The fault feature indexes obtained by different methods KR_s, R_s and K for simulated signals are listed in Table 2. It can be observed that the proposed method in this paper has the larger KR_s, which means that the proposed method is more effective than SK in bearing fault diagnosis under the background of impulse noise.

7 Experimental Validations

In this section, two experimental cases were investigated to verify the effectiveness of the proposed method.

7.1 Case 1: Vibration Signal from Case Western Reserve University

In Case 1, we used the rolling bearing vibration experimental data provided by Case Western Reserve University Bearing Data Center [28–30] to verify the proposed method. Since the data set is a small bearing experiment under laboratory conditions, its impulse noise and other noises are relatively small, which can be used to verify the effectiveness of the algorithm in the case of high signal-to-noise ratio. The inner race fault vibration signal numbered 171 of the drive end bearing is selected which has a defect size of 0.014 in. The sampling frequency of the experimental data is 12 kHz, the data length is 10,000 data points, and the shaft speed is 1750 r/min. According to the theoretical calculation of the rolling bearing theory [31], the inner race fault frequency f_i is 158 Hz.

First, we still use the Fast Kurtogram algorithm to obtain the Kurtogram map as shown in Fig. 5. The optimal frequency band is the red circle in the figure. A

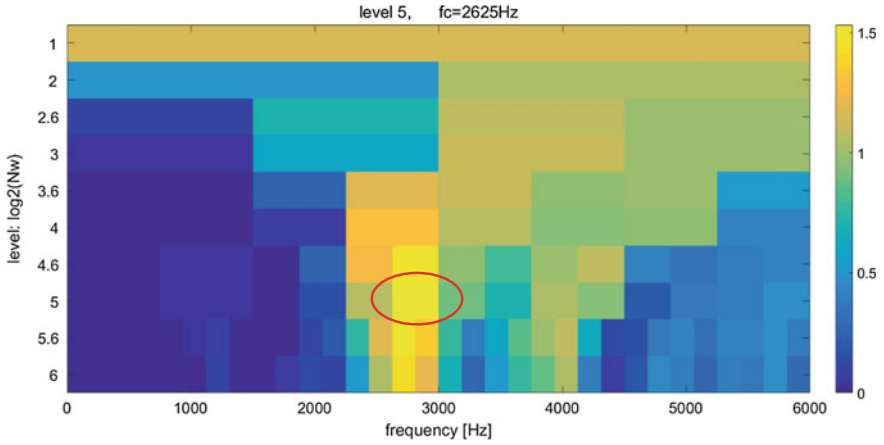


Fig. 5 Fast Kurtogram of 171

filtered signal having a center frequency $f_c = 2968\,\text{Hz}$, a bandwidth $B_w = 156\,\text{Hz}$ is obtained by a band pass filter, and a corresponding square envelope analysis is performed, as shown in Fig. 6. From the figure we can see that the inner race fault is diagnosable, but there are many discrete components which may interfere with the diagnosis.

Corresponding cyclic domain spectrum is shown in Fig. 7. The theoretical fault frequency of the inner race fault can also be easily found from the figure, and the fault frequency and its harmonics are more prominent than those of Fig. 6. From

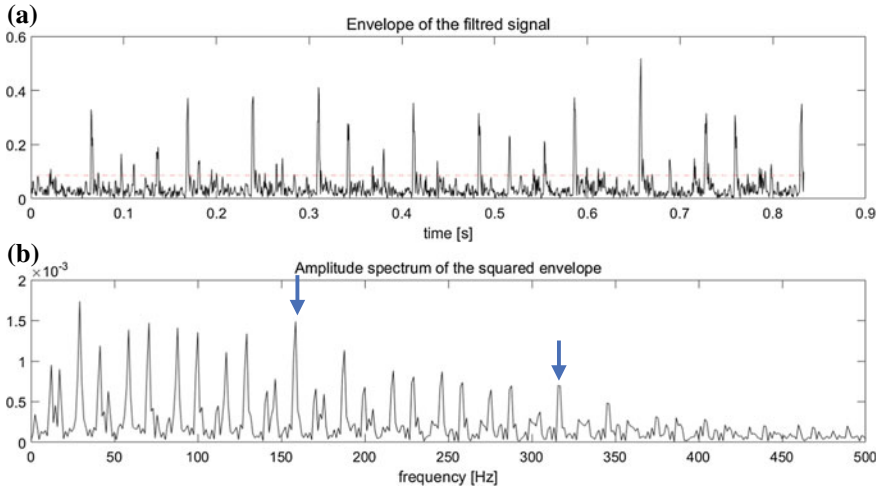


Fig. 6 **a** Envelope of the filtered signal which maximizes the Kurtogram. **b** Amplitude spectrum of the squared envelope

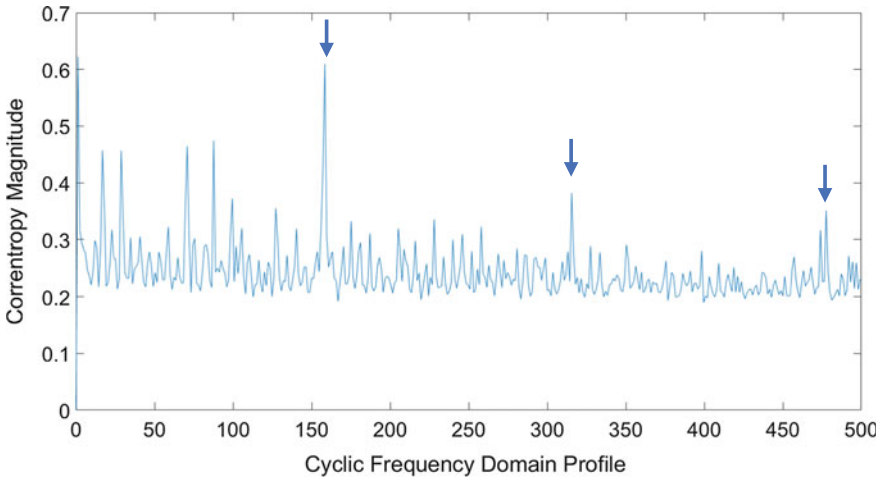


Fig. 7 Cyclic domain profile of inner race fault of 171

Table 3 The fault feature indexes of the inner race fault signal using different methods

171	Proposed method	SK
K	22.455	10.876
R_s	0.013	0.024
$K\,R_s$	0.299	0.264

the comparison the index of the diagnostic effect evaluation in Table 3, the same conclusion can be drawn.

7.2 Case 2: Railway Axle Bearing Fault Data

In Case 2, we will analyze and solve some practical problems in the practical diagnosis of railway bearing fault diagnosis. Due to the complexity of the railway system's operating conditions, a unique component of industrial railway bearing signals that differ from other application scenarios is impulsive noise. During the running of the train, the collision of the wheel with the track, the impulse during the turning process will be transmitted to the bogie via the wheel and the shaft. At this time, there will be some impulsive noise on the bearing, and this phenomenon poses a challenge to the online monitoring fault diagnosis of the bearing. In order to study this problem, as shown in Fig. 8, the axle rolling bearing with real faults collected from the working train was simulated on the train bearing test bench, and the vibration signal under the impulse noise was obtained.

The test bench is driven by a motor. The vertical load simulates the vertical load of the train. The lateral load simulates the contact collision between the wheel and the

Fig. 8 Test rig of railway axle rolling bearings

track during the turning of the train. The vibration sensor is mounted at 12 o'clock (vertical) and 3 o'clock (horizontal) of the bearing case to acquire vibration data. The sampling frequency is set to 12,800 Hz, the running speed is set to 90 km/h, the vertical load is set to 272 kN, and the lateral load which may generate impulse noise is set to 20 kN. A set of inner race fault bearing and outer race fault bearing were tested. According to the transmission ratio of the experimental platform and the bearing structural parameters, the inner race fault frequency and the outer race fault frequency were calculated to be 124 Hz and 89 Hz, respectively.

The algorithm proposed in this paper is used to analyze the outer race fault signal. The cyclic frequency domain spectrum is shown in Fig. 9. There is a distinct peak at the theoretical fault frequency of 89 Hz and its harmonics (at the blue arrow in the

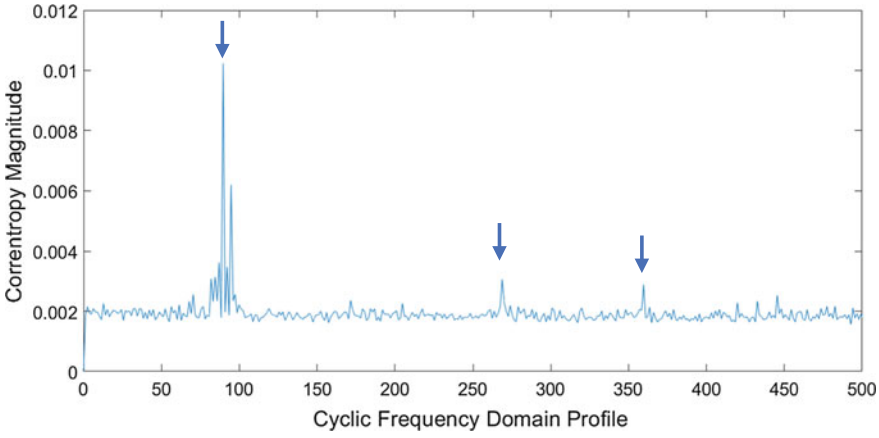


Fig. 9 Cyclic domain profile for railway axle bearing signal of outer race fault

figure). The results show that the proposed algorithm can effectively detect the outer race fault of the rolling bearing. For comparison, the same signal is analyzed using the Fast Kurtogram algorithm. The resulting Kurtogram is shown in Fig. 10. The optimal band is circled in red. A filtered signal having a center frequency $f_c = 6200$ Hz, a bandwidth $B_w = 187.5$ Hz is obtained by a band pass filter, and the corresponding square envelope analysis is performed as shown in Fig. 11. The results show that there is a lot of noise in Fig. 11b, the interference frequency is dominant in the figure, and it is difficult to identify the fault frequency of the outer race. The above

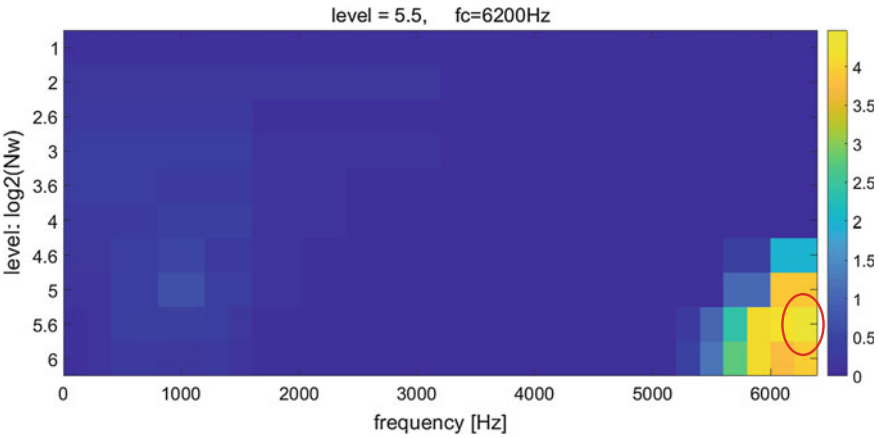


Fig. 10 Fast Kurtogram of railway axle bearing signal of outer race fault

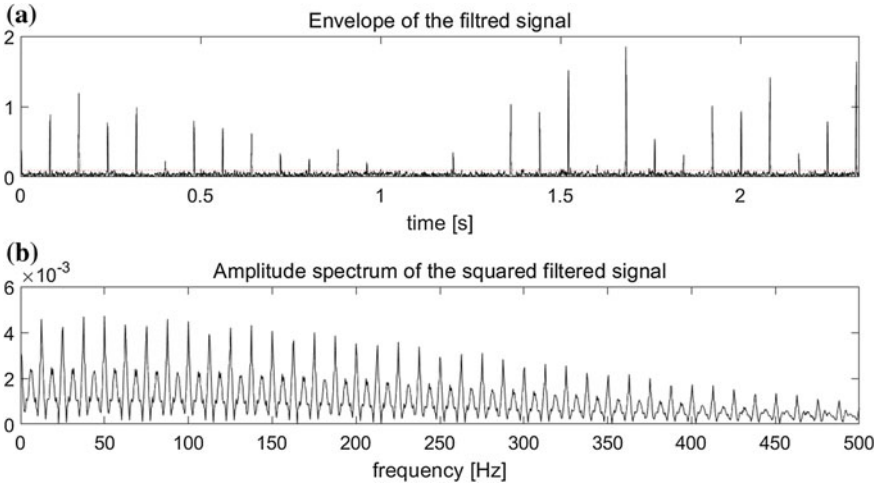


Fig. 11 **a** Envelope of the filtered signal which maximizes the Kurtogram. **b** Amplitude spectrum of the squared envelope

Table 4 The fault feature indexes of the outer race fault signal using different methods

Outer race fault	Proposed method	SK
K	171.141	5.789
R_s	0.027	0.012
$K R_s$	4.594	0.069

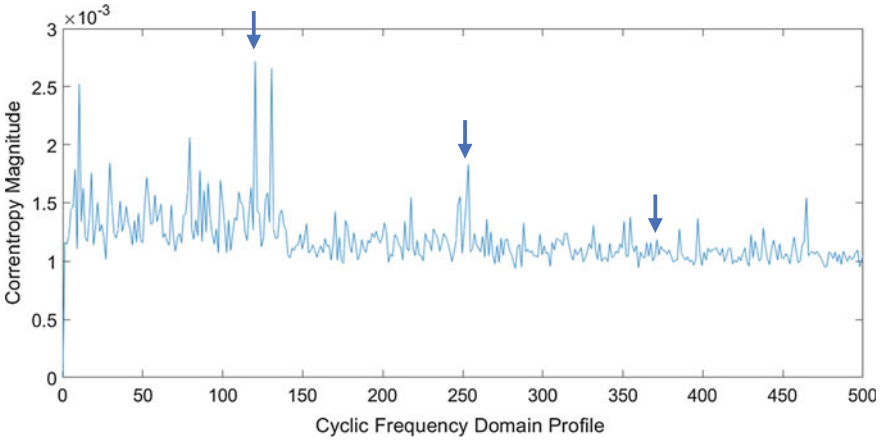


Fig. 12 Cyclic domain profile for railway axle bearing signal of outer race fault

results show that the proposed method has better identification effect on train rolling bearings with impulse noise.

At the same time, the fault feature evaluation index from Table 4 can also quantitatively show that the proposed method has better fault diagnosis effect than the SK algorithm under impulse noise.

The previous procedure is applied to analyze an axle inner race fault signal. The cyclic frequency domain spectrum is shown in Fig. 12. The theoretical fault frequency of 124 Hz and its harmonics can be clearly identified in the figure, as indicated by the blue arrow. The results show that the proposed algorithm can effectively diagnose the inner race fault. The Fast Kurtogram analysis is performed on the same inner race signal, and the Kurtogram is obtained as shown in Fig. 13. The filtered signal of the center frequency $f_c = 6200$ Hz, the bandwidth $B_w = 187.5$ Hz is obtained by the band pass filter, and the corresponding square envelope analysis is performed in Fig. 14.

Unfortunately, in Fig. 14b, the fault frequency information is again submerged in the background noise, and the theoretical inner race fault frequency is difficult to identify. At the same time, according to the fault evaluation index in Table 5, it can be clearly obtained, and the SK algorithm is weak for the analysis of the signal containing the impulse noise. The results show that the proposed algorithm can clearly diagnose the inner race fault signal with impulse noise, and its effect is obviously better than SK algorithm.

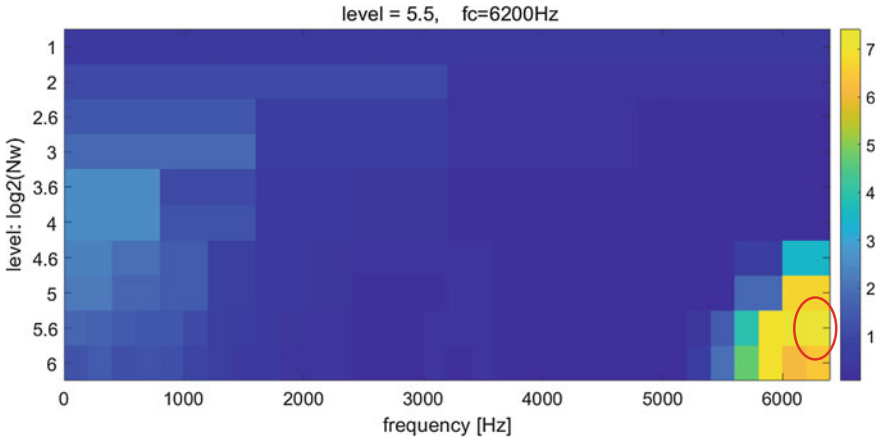


Fig. 13 Fast Kurtogram of railway axle bearing signal of inner race fault

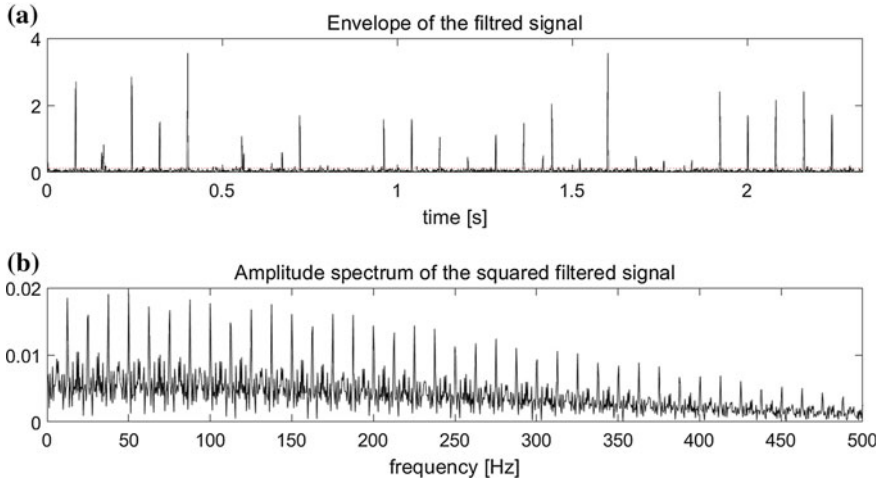


Fig. 14 **a** Envelope of the filtered signal which maximizes the Kurtogram. **b** Amplitude spectrum of the squared envelope

Table 5 The fault feature indexes of the inner race fault signal using different methods

Inner race fault	Proposed method	SK
K	18.054	8.690
R_s	0.016	0.009
$K\,R_s$	0.280	0.079

8 Conclusion

In this paper, we can use the correntropy function to effectively suppress the function of the anomaly point. So, Based on the cyclostationary analysis of EMD and related entropy functions, we propose a fault diagnosis algorithm for rolling bearings that can effectively suppress impulse noise, based on EMD and CCE. The effectiveness of the algorithm is fully verified by using simulated signal and two-sets of real bearing data. In addition, we propose a feature index for evaluating the effectiveness of fault diagnosis, which can be used to quantitatively evaluate the effects of different methods in spectral analysis. The experimental results show that the method has good fault diagnosis effect on rolling bearings, and is not only suitable for small motor bearings, but also for the diagnosis of industrial railway bearings. In particular, it also has a strong identification effect on signals containing impulsive noise, and is significantly superior to the widely used SK algorithm.

Acknowledgements This research is supported by the National Key Research and Development Program of China (Grant No. 2016YFB1200505-014), National Natural Science Foundation of China (Grant No. 61833002).

References

1. Zeng Gang F (2010) Research on detection method of weak fault signal of rolling bearing. Chengdu University of Technology
2. Ming Y, Chen J, Dong G (2011) Weak fault feature extraction of rolling bearing based on cyclic Wiener filter and envelope spectrum. Mech Syst Sig Process 25:1773–1785
3. Randall RB, Antoni J (2011) Rolling bearing diagnostics—a tutorial. Mech Syst Sig Process 25:485–520
4. Antoni J, Randall RB (2006) The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. Mech Syst Sig Process 20:308–331
5. Antoni J (2006) The spectral kurtosis: a useful tool for characterizing non-stationary signals. Mech Syst Sig Process 20:282–307
6. Antoni J (2007) Fast computation of the kurtogram for the detection of transient faults. Mech Syst Sig Process 21:108–124
7. Smith WA, Fan Z, Peng Z, Li H, Randall RB (2016) Optimised Spectral Kurtosis for bearing diagnostics 456 under electromagnetic interference. Mech Syst Sig Process 75:371–394
8. Antoni J (2016) The infogram: Entropic evidence of the signature of repetitive transients. Mech Syst Sig Process 465(74):73–94
9. Gardner WA (1986) The spectral correlation theory of cyclostationary time-series. IEEE Trans Sig Process 11(7):13–36
10. Gardner WA (1987) Spectralcorrelation of modulated signals: PART I-analog modulation. IEEE Trans Commun 35(6):584–594
11. Gardner WA (1991) Exploitation of spectral redundancy in cyclostationary signals. IEEE Trans Sig Process 16(2):14–36
12. Antoni J, Bonnardot F (2004) Cyclostationary modelling of rotating machine vibration signals. Mech Syst Sig Process 18(6):1285–1314
13. Antoni J, Xin G, Hamzaoui N (2017) Fast computation of the spectral correlation. Mech Syst Sig Process 92(474):248–277

14. Gómez-Chova L, Jenssen R, Camps-Valls G (2012) Kernel entropy component analysis for remote sensing image clustering. IEEE Trans Geosci Remote Sens Lett 9(2):312–316
15. Santamaria I, Pokharel PP, Principe JC (2006) Generalized correlation function: definition, properties, and application to blind equalization. IEEE Trans Sig Process 54(6):2187–2197
16. Luan S, Qiu T, Zhu Y (2016) Cyclic correntropy and its spectrum in frequency estimation in the presence of impulsive noise. Sig Process 120:503–508
17. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N-C, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc A Math Phys Eng Sci 454:903–995
18. Zhang J (2010) Comparative study on the performance of commonly used methods for suppressing the end effect of empirical mode decomposition. Yunnan University, Kunming
19. Zhang X, Pan H, Zhang Y (2014) Gearbox fault diagnosis based on particle filter and HHT. Combined Mach Tool Autom Process Technol 71–74
20. Liu B, Yan L, Zhou D (2006) A comparative study of several classical similarity measures. Comput Appl Res 23(11):1–3
21. Wang P, Qiu T, Jin F (2018) A tough DOA estimation method based on sparse representation under impulse noise [J]. Acta Electronica Sinica 46(07):1537–1544
22. Liu W, Pokharel PP, Príncipe JC (2007) Correntropy: Properties and applications in non-Gaussian signal processing. IEEE Trans Sig Process 55:5286–5298
23. Jeong KH, Liu WF, Han S et al (2009) The correntropy MACE filter. Pattern Recognit 42(5):871–885
24. Gunduz A, Principe JC (2009) Correntropy as a novel measure for nonlinearity tests. Sig Process 89(1):14–23
25. Chen B, Xing L, Zhao H, Zheng N, Principe JC (2016) Generalized correntropy for robust adaptive filtering. IEEE Trans Sig Process 64:3376–3387
26. Silverman BW (1986) Density estimation for statistics and data analysis. Chapman & Hall, London, UK
27. Antoni J, Randall R (2003) A stochastic model for simulation and diagnostics of rolling element bearings with localized faults. J Vib Acoust 125:282–289
28. Wang Y, Kang S, Jiang Y et al (2012) Classification of fault location and the degree of performance degradation of a rolling bearing based on an improved hyper-sphere-structured multi-class support vector machine. Mech Syst Sig Process 29:404–414
29. Case Western Reserve University Bearing Data Center Website. http://csegroups.case.edu/-bearingdatacenter/home
30. Loparo K (2003) Bearings Vibration Data Set. Case Western Reserve University, Cleveland, OH, USA
31. Smith Wade A, Randall Robert B (2015) Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study. Mech Syst Sig Process 64–65:100–131

Research on Driving Behavior of Mountain City Passenger Car Drivers Based on GPS Data

Ying Chen and Jin Xu

Abstract In order to study the driving behavior characteristics of urban drivers in mountainous cities. In this paper, a modern data processing technology "GPS" has been used. Through GPS acquisition and comparative analysis method, the travel speed value of six passenger cars in Chongqing, which is collected in two days, is processed, and the effective acceleration value is filtered out. From the point of view of the proportion of sharp acceleration and acute deceleration in the driving process of the driver, the driving acceleration of six vehicle drivers is compared and classified, so the driver behavior characteristics are obtained. Then the driving speed of different drivers on the same road section is compared separately, and the behavior characteristics of different drivers for the same speed limit are summarized. The results show that: (1) The habits of different drivers in the driving process are not the same, the experiment out of three models, "Remain Constant", "Preference Acceleration", "Fast and Slow" type; (2) Because of its properties, in the process of driving, the speed range of passenger car is not large. That is, when driving the passenger car, different drivers will control the driving speed in a more stable range; (3) Different drivers treat so-called speed limit signs differently when crossing the same road, the experiment out of two types, "Complete Follow" and "Appropriate to Follow"; (4) Chongqing belongs to the mountain city, the large number of tunnels and bridges in passenger routes makes drivers more vigilant than other plain areas. In this paper, make a quantitative analysis of driver's driving behavior from the angle of GPS data, which is helpful for the management department to control and supervise the driving behavior such as driver speeding, and provides the basis for the improvement of road infrastructure.

Keywords Passenger traffic · Driver behavior · GPS data · Speed · Mountain city

Y. Chen (✉)
School of Transportation, Chongqing Jiaotong University, Chongqing 400041, China
e-mail: 2318969030@qq.com

J. Xu
Chongqing Key Laboratory of "Human—Vehicle—Road" Cooperation & Safety for Mountain Complex Environment, Chongqing 400041, China
e-mail: yhnl_996699@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_113

1 Introduction

At present, with the rapid development of China's social economy, vehicle parc as well as passenger and cargo transport turnover and road mileage continue to grow, the whole traffic system pressure is increasing. And in the whole safety of "driver-vehicle-road", people play a vital role. In a series of human activities, driving, as a special, complex and dangerous. In order to achieve the correct control of the vehicle, the driver should to maintain a high degree of concentration in the driving process, so that they can quickly and accurately perceive and judge the external environment information [1]. In the past decade, through the analysis of traffic accident data caused by key violations such as speeding, drink driving and illegal overtaking, statistics show that traffic deaths caused by illegal behaviors account for 40% of the total number of deaths. And in recent years, the number of deaths caused by speeding is still on the rise [10]. Therefore, standardizing the driver's driving behavior has become an important guarantee to promote the smooth road and the safety of vehicle driving.

There are currently differences in the statistics of drivers' risky driving behavior in various countries, but the risk driving behavior of motorists is an important cause of traffic accidents, which has been recognized. In 1975, Robertson and others find that the average number of violations recorded by the drivers over the two years was significantly higher than that of motorists who had not had a life-threatening traffic accident. On the basis of Robertson's research, Rajalin found that drivers of non-responsible parties to the accident had a relatively high frequency of driving behavior. In the 1990, Reason constructed DBQ in order to quantify the driver's driving behavior, through research, he divided unsafe driving behavior into Driver error and violation [2–8]. The research on risky driving behavior in China started late, but scholars still obtained good research results. In the 1994, Jin [9] and other people applied the risk perception assessment questionnaire to analyze the driver, and concluded that there were significant differences between the accident group and the non-accident group in the driver, and the conclusion that risk sensing was related to driving experience and personal experience. In the 2013, Zheng Dongpeng used questionnaire method and simulation method to study the risk perception of motorists and its influencing factors. However, comprehensive research is limited by traditional methods of data collection, the content of driving behavior analysis is often limited to the analysis of over speed behavior, while ignoring the impact of rapid acceleration and rapid deceleration and other factors [11, 12].

With the rapid development of social technology, GPS Technology [13] as a global positioning system, has been widely used in modern times. In order to carry out centralized management education and reduce the occurrence of traffic accidents, GPS data can more objectively reflect the driver's driving behavior habits [14–21]. However, due to the accuracy of the required data, the complexity of the environment, although it is very common to use GPS positioning system to judge whether vehicles are speeding, the research and application of driving behavior from the microscopic field is still relatively slow.

And because of its unique topography [22], the main roads in mountain cities are extremely irregular and the line of sight is not wide enough, this makes driving difficult for the most part. Therefore, in dangerous areas the driver will be more cautious, but to the slightly flat terrain of the road will be relaxed vigilance, resulting in traffic accidents. And in the type of vehicle in which the accident occurred, the passenger car because of its passenger transport properties, in the event of a traffic accident, casualties and losses are extremely heavy. Therefore, by collecting GPS data to study the driving behavior of passenger car drivers in mountainous cities, the analysis of driving acceleration, deceleration speed and emergency deceleration behavior of different drivers can understand the behavior habits of vehicle drivers more accurately from the quantitative point of view, detect driving habits, and manage education in a timely manner. And then reduce the occurrence of traffic accidents. As an important part of the field of road safety, it is necessary.

This research mainly starts with the acquisition and acquisition of GPS data in a broad sense, and introduces the processing and calculation method of GPS data. Then through the GPS data collection and collation, analyzed the processing of the data results, the driving behavior habits from different drivers has been researched. This study can cause the traffic management department to monitor, adjust and educate the driver's driving behavior in time. For the same location with high frequency of bad driving habits of different drivers, corresponding prompt measures can be taken to reduce the occurrence of traffic accidents.

2 Experimental Scheme

2.1 Experimental Road

Chongqing, as a typical mountain city, six representative passenger transport direction has been selected as the experimental section. The main factors to be considered in this paper are the road state, that is, the transportation route which can fully reflect the combination of multi-tunnel, bridge and road in mountainous cities. The road through which passenger cars are travelling in different directions is shown in Table 1. Figure 1 is a road map.

2.2 Test Vehicle

The experiment mainly analyzes the driving speed data of passenger cars. The experimental model for the medium and higher passenger vehicles, as shown in Fig. 2.

Table 1 Passenger cars through the road

Number	License plate number	Passenger transport direction	The path taken along the way
1	Yu A09G92	Dazu to Beibei	676 county highway-Yurong expressway (Jiuding mountain tunnels to Jiuyun mountain tunnels)-Yuxi interchange-Chongqing ring expressway-Ring Yuwu interchange-Lu Zuofu way
2	Yu A1G675	Tangba to Chongqing	High way 319-Linjiang street-Xiachen way-Shaoyun interchange-Chengdu-chongqing ring high way (Yunwu mountain tunnels to Shuangbei Jialing river bridge)-Songshi road-Yuao road
3	Yu A1G330	Tongnan to Beibei	Tongnan peijiang bridge-Qiaonan road-Jintong road-Tongnan interchange-Chengdu-chongqing ring high way (Yunwu mountain tunnels)—Chongqing ring expressway
4	Yu A1G262	Tongnan to Chongqing	Tongnan peijiang bridge-Qiaonan road-Jintong road-Tongnan interchange—Chengdu-chongqing ring high way (Yunwu mountain tunnels to Shuangbei Jialing river bridge)—Songshi road—Yuao road-Shangqing temple way
5	Yu A1G202	Dazu to university town	676 county highway—Yurong Expressway-University east road
6	Yu A0G208	Dazu to Bishan	676 county highway—Yurong Expressway-Bishan interchange-Daishan road-Jingshan way-Bitong way

Yurong Expressway passes through four tunnels, namely Jiuding mountain tunnels, Bayue mountain tunnels, Yunwu mountain tunnels and Jinyun mountain tunnels

Chengdu-Chongqing ring high way passes through Yunwu mountain tunnels, Pijia bridge, Lai Jiaqiao highway interchange, Shuangbei tunnels and Shuangbei Jialing river bridge

2.3 Methods of Obtaining Experimental Data

This experiment is to obtain the data value by locating the natural driving vehicle during the driving process, and the main data is the speed and time in the driving process. Under the natural state, the vehicle driving is less affected by the artificial test psychology, so that the driver can drive freely in a normal state, more reflects the normal tendency of the driver to drive. Positioning refers to the GPS receiver by

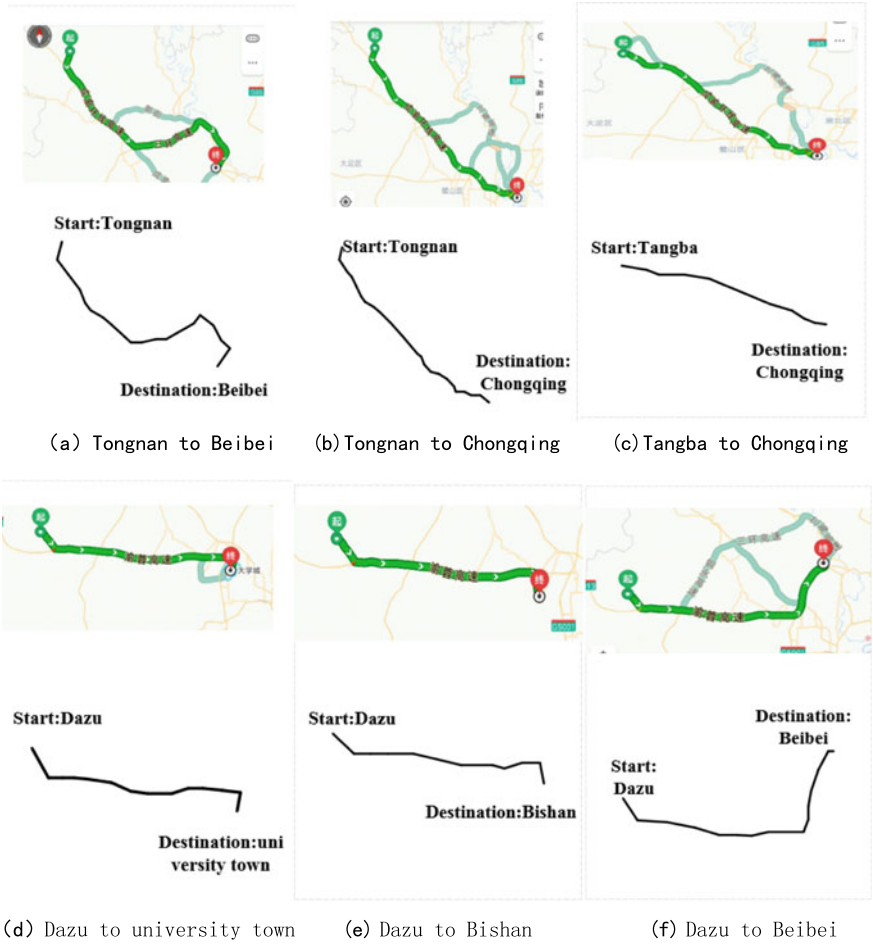


Fig. 1 Road map of test vehicle driving

receiving navigation messages and observation files from the satellite to calculate the location of the current equipment, the basic principle of which is to decode the corresponding satellite signal through the satellite navigation message, and through the calculation of the user's speed, location, posture and other information. The time interval of the ordinary GPS locator is about 1–2 s.

GPS positioning includes static positioning and dynamic positioning. Static positioning refers to the absence of observable changes in the ground-solid coordinates of the points to be measured, so the position coordinates of the points to be measured are determined by means of long continuous observation. Dynamic positioning refers to the observation of a certain observation point, only once, each moment of the point to be measured are different, so dynamic positioning can't converge the positioning results for a long time, so that can used in the real-time positioning.

Fig. 2 Test vehicle

The acquisition of GPS data in this paper is mainly through dynamic positioning, through the Dazu to Beibei, Tangba to Chongqing, Tongnan to Beibei, Tongnan to Chongqing, Dazu to University City and Dazu to Bishan six passenger cars are as the object. Their running speed from September 1, 2018 to 3rd was collected separately and analyzed at intervals of one second.

2.4 Experimental Data Processing and Analysis Theory

For the calculation of data acceleration, the relationship between velocity time and acceleration is mainly obtained, as shown in Formula 1:

$$a = \frac{v}{t} \tag{1}$$

In the calculation Formula (1):

A—accelerated speed (m s^{-2})
V—velocity (km h^{-1})
T—time (s)

The software used for the calculation is Excel and SPSS, after collating the obtained GPS data, by filtering the invalid points, the acceleration of each point is calculated, and the threshold method is further processed. The so-called threshold method refers to the empirical threshold limiting conditions for kinematics information such as acceleration. The judgment of the rapid acceleration and deceleration of driving behavior studied in this paper is to extract the corresponding acceleration information through the original data, and to give the empirical threshold, the discriminating method is relatively simple.

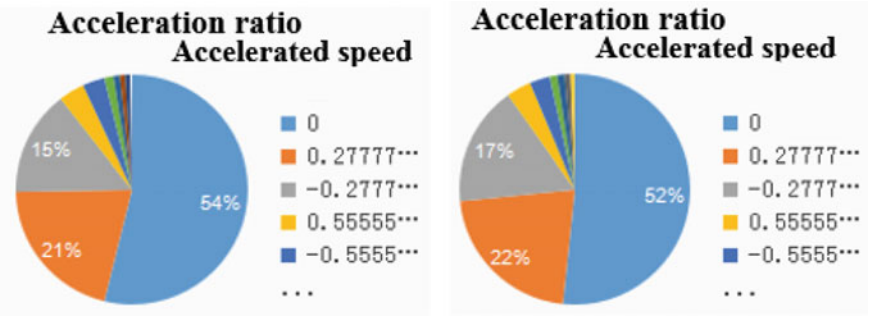
In general, when speed and time are beyond the specified threshold, the sudden change velocity behavior is considered to have occurred. The sudden change speed includes a sharp acceleration and a sharp deceleration, because there is no clear definition of what degree of acceleration and reduction is the true sudden change speed. The threshold given in this paper is also only for this paper. Through the analysis of the distribution of experimental data, this thesis takes $(-3.5, +3.5\ \mathrm{m/s^2})$. When the acceleration exceeds $3.5\ \mathrm{m/s^2}$, it is considered that there is the possibility of rapid acceleration, and when the reduction rate is less than $-3.5\ \mathrm{m/s^2}$, it is considered that there is a possibility of rapid deceleration.

3 Experimental Results

3.1 Driving Acceleration Model of Bus Driver

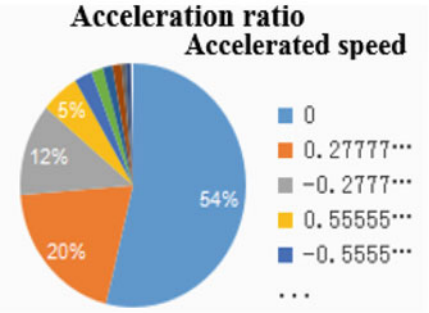
Through the GPS data obtained, the acceleration in the driving process of passenger cars is calculated by adding speed and speed, as shown in Figs. 3, 4, 5, 6, 7 and 8. The picture is the analysis of the acceleration distribution of different passenger vehicles on the corresponding road, and Table 2 is the proportion of emergency acceleration and deceleration of different drivers.

By comparing the acceleration distribution of the same driver in the same vehicle for two days, the overall variation characteristics and distribution characteristics of different drivers can be analyzed. We found that, in addition to Tangba to Chongqing, in the Fig. 6b, the second day of this route of data reception has a certain problem resulting in too little data, distribution and most of the situation is different, the other 5 experimental road passenger car driving acceleration in two days of change is not very large. According to the distribution of acceleration, the driver driving mode can be classified.

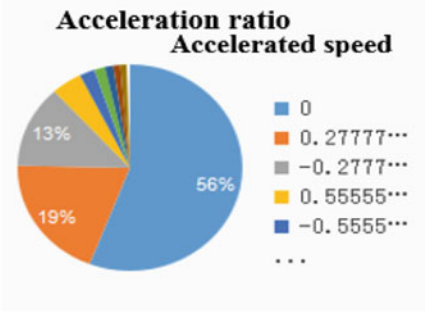


(a) The First Day from Tongnan to Chongqing (b) The Second Day from Tongnan to Chongqing

Fig. 3 Proportion of vehicle acceleration distribution between Tongnan and Chongqing

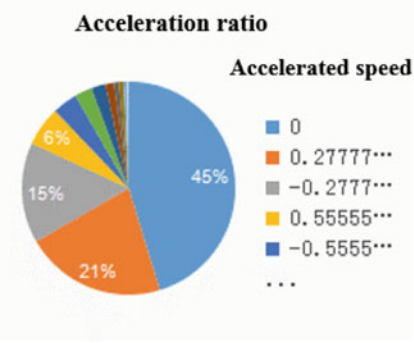


(a) The First Day from Dazu to University Town

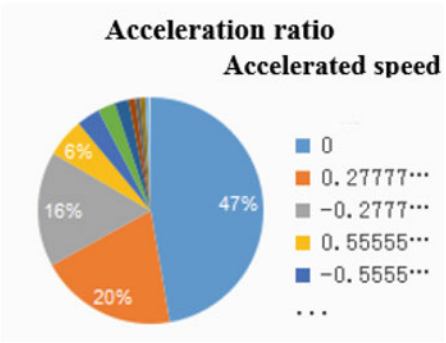


(b) The Second Day from Dazu to University Town

Fig. 4 Proportion of vehicle acceleration distribution between Dazu and university town

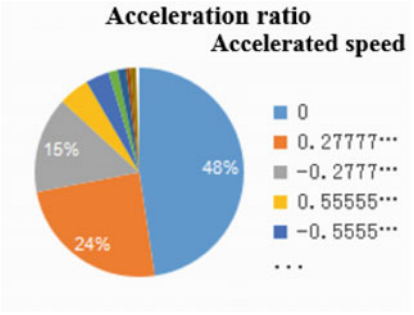


(a) The First Day from Dazu to Beibei

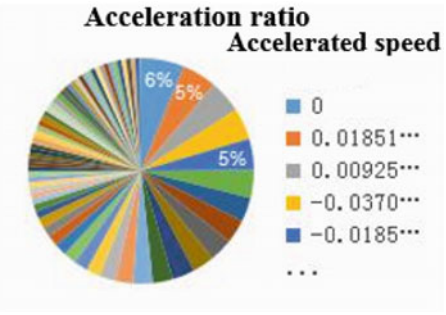


(b) The Second Day from Dazu to Beibei

Fig. 5 Proportion of vehicle acceleration distribution between Dazu and Beibei

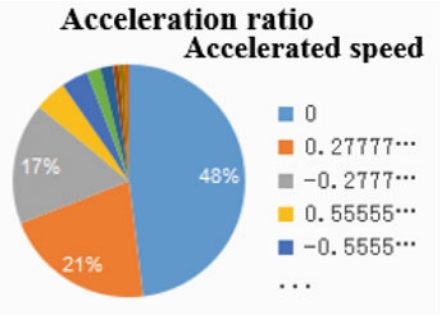


(a) The First Day from Tangba to Chongqing

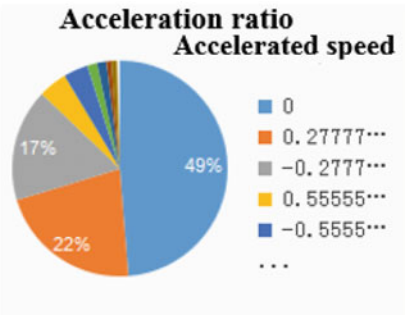


(b) The Second Day from Tangba to Chongqing

Fig. 6 Proportion of vehicle acceleration distribution between Tangba and Chongqing

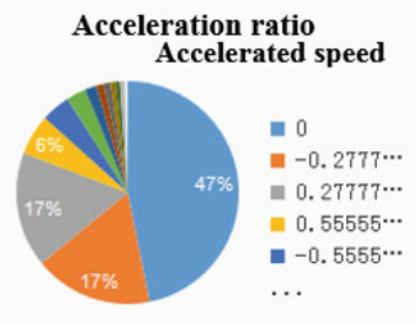


(a) The First Day from Tongnan to Beibei

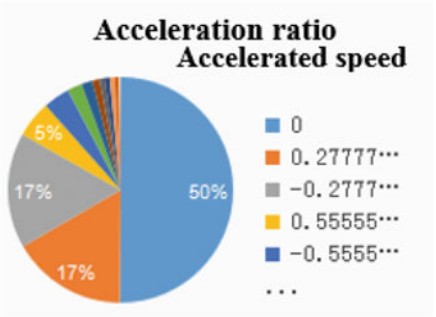


(b) The Second Day from Tongnan to Beibei

Fig. 7 Proportion of vehicle acceleration distribution between Tongnan and Beibei



(a) The First Day from Dazu to Bishan



(b) The Second Day from Dazu to Bishan

Fig. 8 Proportion of vehicle acceleration distribution between Dazu and Bishan

Table 2 Rapid expedited deceleration ratio of vehicles

Number	Vehicle	Rate of rapid acceleration and deceleration
1	Yu A0G208	$2.45 * 10^{-2\%}$
2	Yu A1G202	$0.56 * 10^{-2\%}$
3	Yu A1G262	$1.8 * 10^{-2\%}$
4	Yu A1G303	$1.4 * 10^{-2\%}$
5	Yu A1G675	$0.19 * 10^{-2\%}$
6	Yu A09G92	$0.41 * 10^{-2\%}$

The first type of pattern: the "Remain Constant" driver, as shown in Figs. 3a, b, 4a, b and 8b.

Tongnan to Chongqing, Dazu to the university town, Dazu to Bishan passenger car in the course of driving, the acceleration of zero reached 50% and above, marking the driver in the driving process is basically maintained in a more uniform range. When

the driver is driving through the tunnel, the speed will be properly slowed down. Once leaving the tunnel, they will be corresponding acceleration. But the acceleration is still maintained in a relatively small range, subjectively think that there will be no urgent and urgent reduction of the situation, but we can see from the analysis of the rapid acceleration and deceleration that in this mode the driver's rapid acceleration and rapid deceleration is in the medium level.

The second type of pattern: The driver of "Preference Acceleration", as show in Figs. 5a, b and 6a.

It can be seen from the acceleration map that this type of driver accounts for a larger proportion of acceleration compared to deceleration during driving.

Drivers in this mode tend to be more edgy in character, and they will continue to accelerate when the road conditions is better. The biggest drawback is that if there is a problem with the road conditions, it will slow down at a greater rate of reduction. Subjective drivers in this mode are more likely than other models to experience frequent and rapid deceleration of driving habits. However, it can be seen in the rapid acceleration and deceleration ratio analysis of the driver's driving data that the proportion of drivers in this category is not very high.

The third type of mode, "Fast and Slow" drivers, as shown in Figs. 7a, b and 8a, b.

We can see from the Tongnan to Beibei and Dazu to Bishan that, in addition to maintaining a uniform speed during the driving process, the driving process is always in an unstable state that accelerates and slows down for a while. The proportion of the same acceleration and deceleration on the graph is almost identical, the driver is in a tangled state, on the one hand, want to speed up, on the other hand, fear of danger, and reduce the speed to the original level. And so on and so forth. In this mode, the driver in the driving process constantly change the speed to maintain at a level, this will make the driving process more tense, more likely to cause fatigue driving. Subjectively, the driver of this situation generally does not have a sharp acceleration of the rapid deceleration situation, but through the integration of Table 2 and Figs. 3, 4, 5, 6, 7 and 8, it can be seen that the third model of the driver preference for "Fast and Slow", drivers in this mode accelerate and decelerate more rapidly than drivers in the other two modes.

In general, the "Preference Acceleration" of the driver in the driving process at a high level of speed, the probability of rapid acceleration and deceleration appears the lowest; The "Remain Constant" driver's driving process is mostly uniform, but once the uniform level has been broken, it is likely to be caused the rapid acceleration and deceleration; The "Fast and Slow" driver often changes his driving speed during the driving process, and from the speed distribution its seems to control the driving speed in a stable level, but because of its rapid acceleration and deceleration ratio is highest, it can be inferred that such drivers is not focused enough in the driving process.

Therefore, compared with other two types of drivers, the "Fast and Slow" driver are slower to respond to emergencies, and more likely to lead to traffic accidents.

3.2 *Analysis of the Critical Value of Driving Speed*

In the process of speed data analysis of passenger cars in six directions, it can be seen that the driving route in six directions will have a certain coincident section. However, on these coincident sections, different drivers will adopt different coping options for the same road indication requirements. As shown in Fig. 9, using SPSS software to intercept the partial velocity distribution of six vehicles under continuous sections with a speed limit of 60–80–90–100 km/h can be broadly divided into the following two categories:

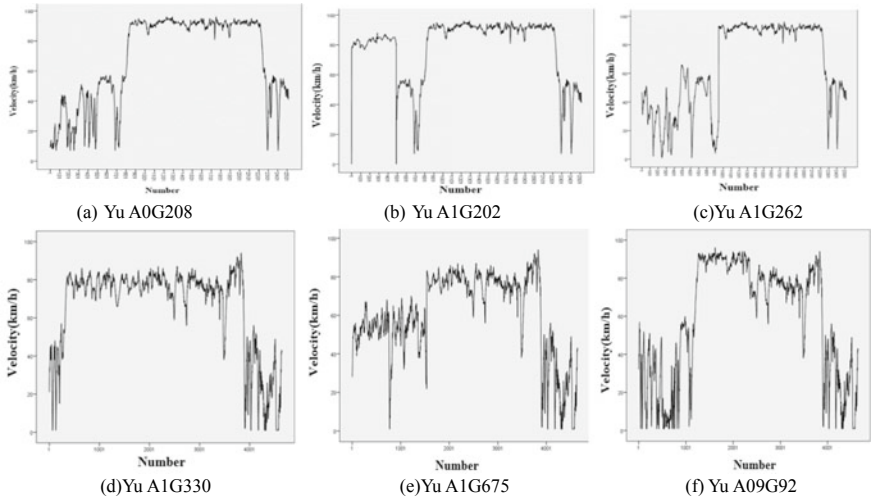
Category I: Decelerate completely below the speed limit according to the speed limit sign.

As shown in Fig. 10, this type of driver belongs to the safety driver. In the process of driving, the operator will strictly follow the road instructions and the driving speed value will also maintained at a stable level. So, there will be no big mistakes in the driving process.

Category II: Slow down to critical speed.

Such drivers will decide, depending on the situation, whether or not to follow the instructions. In this case, the velocity data shown in Table 3. When the speed limit is 60 km/h, the driver will not be completely slowed down below the speed limit, but will control the driving speed within the approximate range of the speed limit value.

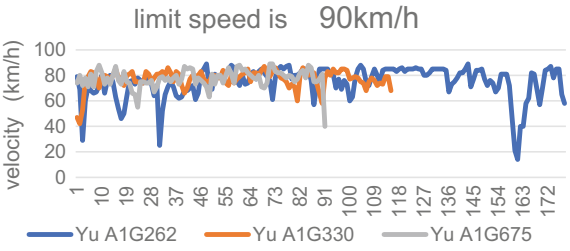
In the analysis of the driving speed at the speed limit, six passenger car drivers showed the same tendency, that is, when the speed limit is 60 km/h, the driver generally keeps the driving speed close to 60 km/h and below. However, when the speed limit is large, the driver tends to be more inclined to slow down slightly around the speed.



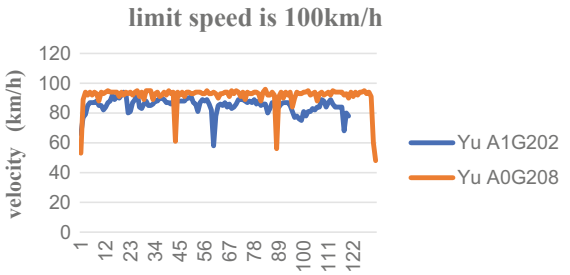
(a) Yu A0G208 (b) Yu A1G202 (c)Yu A1G262

(d)Yu A1G330 (e)Yu A1G675 (f) Yu A09G92

Fig. 9 Speed distribution of speed limit area



(a) The Speed of Different Vehicles is Limited to 90km/h



(b) The Speed of Different Vehicles is Limited to 100km/h

Fig. 10 The speed of different vehicles at the same speed limit

Table 3 Appropriate deceleration at near critical speed

Vehicle	Time	Speed/(km h^{-1})	Rate-limiting/(km h^{-1})
Yu A1G675	2018-8-31	61	60
Yu A1G675	2018-8-31	62	60
Yu A1G675	2018-8-31	61	60
Yu A1G675	2018-8-31	62	60
Yu A1G675	2018-8-31	62	60
Yu A1G675	2018-8-31	61	60
Yu A1G675	2018-8-31	61	60
Yu A1G675	2018-8-31	61	60

4 Conclusion

As an important issue in today's society, the impact of road traffic safety on social and economic development should not be underestimated. As the core factor affecting road traffic, the driver's driving behavior must be paid attention to, so how to effectively standardize the driver's driving behavior and reduce the bad driving habits is very important.

In this paper, GPS positioning is mainly used to obtain the relevant speed data of drivers in natural driving, and through the calculation of its acceleration, the ratio of the driver's sharp acceleration to the sharp deceleration is considered, and then

the driving habit of the corresponding driver is judged. The main conclusions are as follows:

(1) Due to personality, driving age and other internal factors, different drivers in the course of driving will also show different driving habits, through the analysis of the relevant data in this experiment, the driving habits of different drivers are classified, summed up three models, "Remain Constant", "Preference Acceleration", "Fast and Slow" type;
(2) Because of its passenger transport properties, in the course of driving, in order to ensure the comfort of passengers and other reasons, passenger car drivers will control the speeds in a small range. Even for different passenger car drivers, they will both control the speed in a more stable range when driving passenger cars;
(3) Due to the degree of concentration in driving, road conditions and other reasons, in the same road section, different drivers will have different attitude towards the so-called speed limit sign. This experiment divides it into two types, "Decelerate completely below the speed limit according to the speed limit sign." and "Slow down to critical speed";
(4) Chongqing belongs to the mountainous city, the passenger route will appear a large number of tunnels, bridges. Although the continuous and changeable road environment will make the driver compared with other plains areas more vigilant, it cannot be ruled out that the driver's alertness will be reduced after he is fully familiar with the driving route and his concentration will be reduced during the driving.

References

1. Li J (2018) Driver risk driving behavior analysis and related factors. Heilongjiang Sci 9(10):120–121
2. Francesco B (2013) Driver performance approaching and departing curves: driving simulator study. Traffic Injury Prevention 15(3):310–318
3. Bella F (2005) The evaluation of design consistency: predicting models of operating speed on three-dimensional alignment from tests on driving simulator. In: 3rd international symposium on high-way geometric design
4. Antonson H, Ahlström C, Wiklund M et al (2013) Crash barriers and driver behavior: a simulator study. Traffic Injury 14(8):874–880
5. Bella F (2007) Parameters for evaluation of speed differential: contribution using driving simulator. Transp Res 20(23):37–43
6. Bella F, D'Agostini G (2010) Combined effect of traffic and geometrics on rear-end collision risk: a driving simulator study. Transp Res 21(65):96–103
7. Naatanen R, Summala H (1976) Road user behavior and traffic accident. Oxford, North-Holland
8. Van der Horst AR, Hogema JH (2011) Driving simulator research on safe highway design and operation. Transp Res 22(48):87–89
9. Jin H, He C, Chen D (1994) Hazard susceptibility. Psychol Sci 17(1):42–46
10. Yang M, Wang X, Zhu M (2017) Study on driving behavior based on natural driving experiment. Transp Transp 3:7–9

11. Zhao M, Yang Y, Pan X (2010) Experimental study on radial acceleration of flat curve section of two-lane highway. Traffic Sci Eng 26(4):7–10
12. Ding H, Guo K, Li F et al (2010) Arbitrary road and speed follow control driver model based on acceleration feedback. J Mech Eng 46(10):116–120
13. Li Y (2018) On error and precision control analysis of GPS measurement. Smart city 5(38):52
14. Liu Y, Yu Z, Wang Y et al (1993) Global positioning system (GPS) and its applications. Surveying and Mapping Press, Beijing
15. Guo Y (2017) Analysis of risk driving behavior of operating passenger cars based on GPS trajectory data. Edu Teach Forum 3:76–77
16. Yang Y, Yao E, Pan L et al (2015) Research on taxi path selection behavior based on GPS data. Trans Syst Eng Inf 15(1):81–86
17. Yang L (2018) Detection of abnormal behavior of expressway vehicles based on GPS data. J Chongqing Jiaotong Univ (Natural Science Edition) 37(5):97–103
18. Chen K, He K, Shang J et al (2017) Research on detection technology of abnormal driving behavior of expressway based on mobile intelligent terminal. Sci Technol Innov Appl 9:20–21
19. Zhang Y (2018) Driving behavior recognition based on multi-sensor data fusion in mobile terminals. East China Normal University
20. Li H (2004) Investigating morning commute route choice behavior using global positioning systems and multi-day travel data. Georgia Institute of Technology, Atlanta
21. Zuriaga AM, Garcia A, Camacho Torregrosa FJ et al (2010) Modeling operating speed and deceleration on two-lane rural roads with global positioning system data. Transp Res 21(71):11–20
22. Hu L, Jiang Q, Tang M (2010) Study on speed limit of mountainous expressway based on driving behavior. Traffic Sci Eng 26(3):82–88

A Comprehensive Collision Prevention Approach for Rural Highway in Mountain Area

Fengchun Han, Dan Zhao, Wen Shen and Sheqiang Ma

Abstract Many serious traffic collisions occur at rural highways in mountain area during the rapid urbanization in China, when traffic demands are large, vehicle composition is complicated. This brings tremendous pressure on transportation agencies. In order to improve safety performance of rural highways in mountain area of China, a comprehensive collision prevention approach was developed systematically. A case study regarding the effect of before-and-after implementation of this advanced collision prevention approach at Longyan, Fujian Province was conducted. It showed that highway risk level (HRL) evaluated through highway risk assessment method in the technical guide for the implementation of highway safety and life protection engineering was reduced from 19.09 to 7.59 with safety level improved from grade IV to III. The merit of this newly developed approach has been verified and is very promising for future large scale promotion and implementation.

Keywords Rural highway · Collision prevention · Traffic safety · Traffic facility · Risk assessment

1 Introduction

With strong growth of urban and rural economy and tourism in China, the rural highway traffic volume increases very fast. Due to the complex road section conditions and vehicle composition in villages and towns, wide differential in running speed, poor road surface conditions, inappropriate traffic safety facilities and weak awareness of traffic safety of villagers and other reasons, the road sections in villages and towns have many potential challenges of traffic safety, frequent traffic accidents and serious consequences. Therefore, it is of great practical significance to study

F. Han · D. Zhao (✉) · W. Shen · S. Ma
School of Traffic Management, People's Public Security University of China, Beijing, ST Huangyi, China
e-mail: zhaodan@ppsuc.edu.cn

F. Han
e-mail: hfc1966@163.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*, Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_114

the traffic operation characteristics and traffic accident characteristics of rural highway, set up traffic safety and management facilities scientifically, and improve traffic conditions and traffic safety level.

Domestic and foreign scholars have studied the establishment of highway traffic safety facilities. Early Danish studies have shown that effective control of transit highways can reduce injuries and serious injury accidents by an average level about 33%; according to Webster's analysis in the UK, when speed bumps are set on the road, the 85% speed will be reduced by 16 km/h, the number of traffic accidents will be reduced by 71%, and the number of road accidents will be reduced by 8%. In places where there are more bicycles and pedestrians but no pedestrian isolation facilities, the posted speed limit of 30 km/h is recommended [1]. Considering current situation of traffic safety in the road section crossing villages and towns, Yuchi Guan proposed the method of setting up traffic safety facilities in the road section crossing villages and towns, and verified the feasibility of setting up traffic safety facilities with examples [2]. Jinlong Zhao explored the methods to setup highway traffic signs in mountainous areas, and traffic signs on traffic accident hot spots such as long downhill grades, curves and areas influenced by severe weather [3]. Chengcheng Tang and Yong He put forward the installing methods and requirements for other highway traffic facilities, such as roadside subgrade guardrail, bridge guardrails, traffic signs, traffic line markings, line-of-sight guidance facilities, collision MATS, and hedge lanes [4].

Based on the investigation and analysis of the traffic patterns and traffic accident characteristics of typical sections of roads at Longyuan, Fujian Province, a comprehensive improvement approach is developed for the traffic safety promotion of rural highway, and a risk assessment method is conducted through the precise installation of traffic safety management facilities, in order to create a good environment, reduce traffic accidents, and improve rural highway safety level.

2 Rural Highway Traffic Characteristics

2.1 Road Environment Characteristics

Road environment characteristics Longyan city, in Fujian province, is located in mountainous and heavy hilly area, where the terrain and landform of roads is complex, and highway classifications are at Level three or below. The alignments of roads in villages and towns are mostly composed of horizontal and vertical segments, with curve sections accounting for a large proportion. There are many vertical slopes, of which the slope degrees are usually greater than 3%, while commonly with horizontal curves on the bottom of downhill sections.

The driver sight distance issue The buildings on both sides of the road section in villages and towns are dense and close to the roadway, blocking sight distance of drivers. Especially, the study section is located in the area with poor combination of

horizontal and vertical alignments, so the sight distance of drivers is difficult to meet the driving requirements. It is a typical section with frequent accidents.

The road conditions Due to the long-time use of road sections in villages and towns without necessary maintenance, and the fact that commercial vehicles are mostly overloaded, pavements of several road sections have been seriously damaged: the road surface flatness is still worse, therefore it is easily to gather water on the road surface during rainy days, affecting the driving safety.

Traffic facility characteristics According to the investigation, warning signs, prohibition signs and instruction signs in village and town sections in Longyan area are mostly missing and the positions, layout forms and contents are inconsistent with the provisions and requirements of the national standard of Road Traffic Signs and Marking (GB5768-2009); pavement marking, road side guardrail and deceleration and other safety facilities are seriously damaged. Inadequate and inappropriate traffic safety facilities cannot correctly guide the operation of vehicles and protect the vehicles out of control, which reduces the traffic safety performance of this type of roadway sections.

2.2 Traffic Flow Characteristics

Traffic volume and speed characteristics Taking the roadway sections in mountainous areas of Longyan city, Fujian province as the sample of investigation, field data of February 5 and 10, 2018 were collected to obtain the daily traffic volume and running speed data of the road sections in mountainous areas involving provincial roads and county roads, as shown in Tables 1 and 2.

The daily traffic volume of provincial highways is larger than that of county highways, and the daily traffic volume of passenger cars of provincial highways and county highways accounts for 78% of the total traffic flow, followed by motorcycle 12% and truck 10%, but the traffic volume of provincial highway trucks is about 5.4 times of that of county highways.

Table 1 A statistical table of traffic volume and running speed of motor vehicles in mountain villages and towns (provincial roads)

Vehicle type	Daily traffic (veh/d)	≥80 km/h the number of vehicles (veh)	60–80 km/h the number of vehicles (veh)	40–60 km/h the number of vehicles (veh)	≤40 km/h the number of vehicles (veh)
Passenger car	2200	65	1055	805	175
Van	435	25	45	245	120
Motorcycle	310	0	15	80	215
Total	2945	90	1115	1130	510

Table 2 A statistical table of traffic volume and running speed of motor vehicles in mountainous villages and towns (county roads)

Vehicle type	Daily traffic (veh/d)	≥80 km/h the number of vehicles (veh)	60–80 km/h the number of vehicles (veh)	40–60 km/h the number of vehicles (veh)	≤40 km/h the number of vehicles (veh)
Passenger car	1545	25	305	1075	140
Van	80	0	5	50	25
Motorcycle	265	0	15	60	190
Total	1890	25	325	1185	355

The operational speed of provincial roads is higher than that of county roads. Among provincial highways, the speed exceeding 60 km/h accounts for 41.3%, while county highways only account for 18.5%. The speed of passenger cars passing through villages and towns is mainly concentrated among the range of 40–80 km/h, the speed of freight cars is mainly concentrated in the range of 40–60 km/h, the motorcycle speed less than 40 km/h accounts for 72%, but the phenomenon of illegal manned is serious.

Pedestrian crossing characteristics Survey data shows that the walking time for pedestrians to cross street is generally between 20 and 30 s, with an average endurance time of about 26 s and a maximum endurance time of 35 s. The proportion of crossing the street on crosswalk is up to 80%. Due to the narrow width of the road, pedestrians walk across the street at a unique pace.

3 Analysis of Traffic Safety Along Highway Sections Through Villages and Towns

3.1 Traffic Safety Overview

The number of traffic accidents above the general level occurred in the road sections of villages and towns in mountainous areas of Longyan city in Fujian province from 2015 to 2017 is shown in Table 3.

From 2015 to 2017, a total number of 241 traffic accidents occurred in longyan road sections, including 112 people killed and 207 injured. The number of traffic accidents is decreasing yearly, but the decreasing ratio is small. The number of traffic accidents in provincial highways and county highways is close to that in villages and towns, while the number of traffic accidents in township roads ranks at the lowest level.

Table 3 Statistical table of traffic accidents in villages and towns in mountainous areas of Longyan city

Number of accidents				
Year	Highway administrative classification			The total number of accidents
	Provincial road	County road	Township road	
2015	36	39	8	83
2016	35	36	10	81
2017	33	36	8	77
Total	104	111	26	241

3.2 Traffic Accident Characteristics [5]

Traffic accident time characteristics According to the statistics, the traffic accidents of highway sections in villages and towns in the mountainous area of Longyan city in Fujian province happens almost every month of each year.

Traffic accidents are distributed through four seasons of a year with percentages as 27% in autumn, 30% in winter, 27% in spring and 16% in summer. It can be seen that the traffic accidents in winter are the most, and 84% in spring and autumn and winter. Traffic accidents are related to the flow of people and the traffic environment.

As for the traffic accident distribution of 24 h a day, the three time periods of 4:00–8:00 in the morning, 12:00–16:00 in the noon and 16:00–20:00 in the afternoon are the accident prone time periods, accounting for 18%, 24% and 28% respectively. These three periods account for 70% of all day long traffic accidents. The traffic accidents in these three periods are related to the driving environment and the driver's biorhythm.

Traffic accident space characteristics According to the statistics, it is shown that the traffic accidents on ordinary sections accounts for 60%, whereas 13% happens at sharp curves, 12% happens on entrances and exits of villages and towns, 8% happens on long slopes, and 7% happens on continuous curves. Therefore, it is necessary to strengthen the research on traffic operation characteristics and traffic safety facilities of ordinary sections, acute bends, entrances and exits of villages and towns where accidents occurred frequently.

Traffic accident pattern characteristics The patterns of traffic accidents mainly include rear-end, side-swipe, heads-on and right-angle. Collision accidents account for 40% of the total, far exceeding the proportion of other accident patterns. It is the majority pattern of accidents. Turnover accidents account for 21%. Crash accidents account for 17%, higher than the national average. Rolling and side-swipe accidents account for 7% and 5% respectively, and others account for 11%. Among them, collision, rollover and crash accidents account for 77%, and it is caused by the road alignment and surrounding environment of highway sections in villages and towns in the mountainous area.

Traffic accident vehicle type characteristics Commercial vehicle (CV) is the vehicle with the highest proportion of traffic accidents. Statistics show that CV demand accounts for less than 15% of the total flow, however, 33% of traffic accidents

are related to trucks. In addition, 25% of accidents are caused by motorcycles, 20% are caused by agricultural vehicles and 14% are caused by passenger cars. Truck, agricultural vehicle and motorcycle accidents account for 78% of the total, which should be paid great attention by the management department.

3.3 Analysis of the Influencing Factors of Traffic Accidents

Driver Factors The traffic accidents caused by people's misconduct account for 80–90% of the total. According to the investigation, among the traffic accidents caused by drivers' driving behaviors in mountainous villages and towns, insufficient space headway accounts for 29%, over-speed driving accounts for 23%, driving against the operating instruction accounts for 14%, illegal parking accounts for 12%, illegal overtaking accounts for 9%, negligence accounts for 7%, and other 6%. Therefore, the enforcement on drivers must be strengthened to reduce the number of traffic accidents caused by improper driving behaviors.

Vehicle Factors Vehicle type and condition is one of the influencing factors of traffic accidents. There are a large number of vehicles that do not meet safety requirements in rural highways, such as agricultural vehicles, scrapped vehicles and modified vehicles. Traffic accidents can be easily caused by brake failure, steering failure, tire blowout, vehicle fire lighting failure and others, accounting for 40%, 20%, 14%, 13%, 7% and 4% respectively. Among them, brake failure, steering failure and tire blowout accidents account for 74% of the total. Therefore, it is very important to strengthen vehicle performance testing and improve vehicle safety performance.

Road Factors The horizontal and vertical alignment of this section of road is very poor. It mainly includes horizontal and vertical curves. Moreover, the road passes through villages and towns, which are close to the highway shoulder with insufficient side clearance.

Traffic Facilities Factors There are some problems with traffic safety and management devices, such as not installing collision barriers or retaining walls on both sides of the road, no warning post signs at the entrance, insufficient deceleration facilities, lack of traffic signs and markings, traffic signs being blocked, insufficient illumination and so on. Statistics show that more than 50% of traffic accidents are caused by the lack of collision barriers or retaining walls and the absence of emergency lanes and deceleration facilities.

4 Traffic Safety Promotion Technology of Road Sections in Villages and Towns

Rural roads are often with one side or both sides of highway long straight sections, long downhill sections or sharp downhill slopes. The commonly-used traffic safety

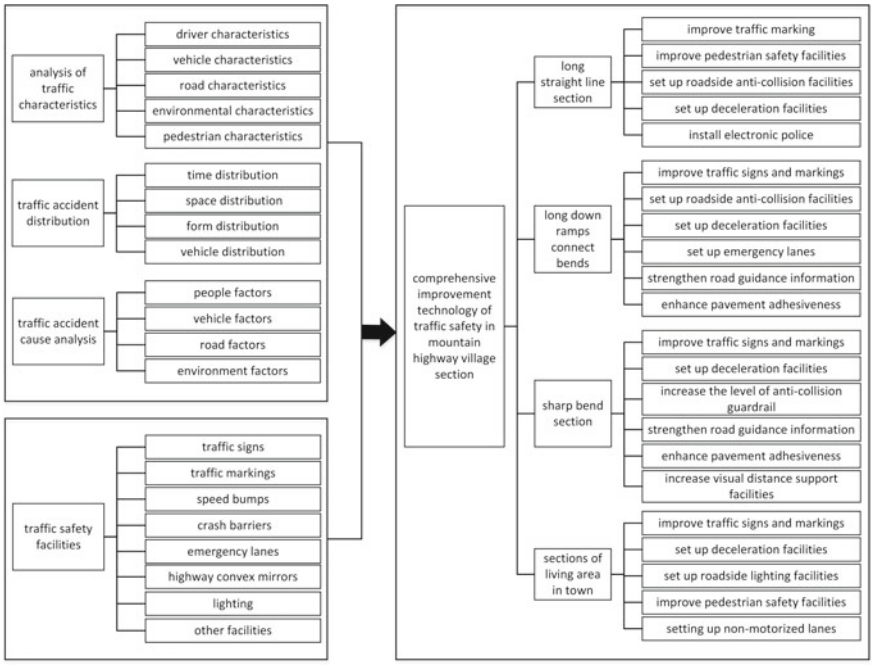


Fig. 1 Framework of comprehensive promotion technology of traffic safety in mountain villages and towns

facilities include: traffic signs, pavement marking, lateral deceleration facilities, roadside guardrail, emergency lanes, retaining walls, visual guidance, speeding capture systems, colored pavement, yellow flashing warning lights and lighting, etc.

Based on the analysis of traffic characteristics, traffic accident characteristics and traffic accident causes of the road sections in mountain villages and towns, and the study on the characteristics of traffic safety facilities, a comprehensive promotion approach of road traffic safety in mountain villages and towns is proposed as shown in Fig. 1.

5 Application Example

5.1 General Situation

The S203 highway in Tieshan Town, Longyan City, Fujian Province, is a long downhill slope section with a grade as 3%, single carriageway, two-way two-lane. The subgrade width is 11 m, lane width is 5 m, and lateral clearance is 0.5 m. There are no isolation facilities in the whole section. There is mountain vegetation on both

sides of the road other than commercial community. The average daily traffic flow of motor vehicles is 2550 veh, and the average speed is 56.7 km/h.

5.2 Traffic Safety Facilities Status

- There are only a few warning signs in this section, and the traffic signs are seriously missing;
- Only the middle lines, the deceleration vibration lines, and road nails wear seriously;
- There is no different grade of crash barriers or retaining walls on the sides of the road;
- Much vegetation blocks traffic signs, which affects the sight distance of drivers.

5.3 The Promotion Technology of Traffic Safety

The main reasons causing traffic safety issues in Tieshan town section of the S203 highway are illustrated as follows: long downhill bend combination alignment, the high speed of the traffic in the sharp bend section, both sides of the road lack of necessary traffic safety management and protection facilities. Combined with the characteristics of traffic flow and traffic accidents of the road sections in mountain villages and towns, the comprehensive promotion technology of traffic safety is proposed as follows:

- **Traffic Signs**

Caution signs: signs that are used for sharp turning and accident-prone sections should be set before corners. Slow-moving signs and village signs should be set before entering villages, warning signs for pedestrians and children should be set before pedestrian crossing (the bottom color of the sign is fluorescent yellow-green), and intersection signs should be set before access points in villages and towns.

Prohibition signs: no overtaking sign, no turning sign and 30 km/h speed limit sign should be set before the bend; stop-controlled sign and no parking sign should be set before the village entrance.

Indication signs: include pedestrian crossing signs.

Directional signs: directional signs should be set before entering the village.

- **Traffic Markings**

Indication lines: include crosswalk and diamond lines, roadway edge white solid vibration lines and reflective road speed limit marking lines.

Prohibition lines: the demarcation lines for forbiting crossing the opposite carriageway are yellow single solid lines, setting stop-controlled lines before entrances of villages and towns.

Caution lines: transverse deceleration vibration lines should be set before and after the straight and curve section; longitudinal deceleration lines should be set at the curve.

- **Other Traffic Safety Facilities**

Transverse cement speed bumps should be set at the entrances of villages and towns.

Color pavement should be set at the corners to enhance the adhesion and visual impact.

Yellow flashing warning lights should be set in front of the crosswalks near the villages and towns.

Reflective track pins should be installed in the center of the road and lighting facilities should be installed along the road.

- **Traffic Control Facilities**

Speed capture equipment should be set up in the straight road sections of villages and towns.

Traffic monitoring equipment should be set up around corners, entrances and crosswalks.

5.4 Risk Evaluation

In order to evaluate the traffic safety improvement levels of Tieshan Town Section of S203 highway, the highway risk assessment method in the Technical Guide for the Implementation of Highway Safety and Life Protection Engineering (Trial Implementation) [6] was used and the result showed that the highway risk (HR) before implementing safety approaches is 19.09, and that after improvement is 7.59. Before the improvement, the risk level of this section is graded IV, with high risk and high probability of traffic accidents. After the improvement, the risk level is graded as III, the medium level. The probability of traffic accidents is reduced, and the level of traffic safety is improved.

6 Conclusion

Rural highway at mountain area plays an important role in rural economic development and new urbanization [7]. Due to the problems of road alignment, road surface condition, traffic volume and traffic safety facilities of the sections, traffic accidents occur frequently and the consequences are serious, which has drawn great attention of relevant agencies. Based on the factual road conditions of the highway sections in villages and towns in the mountainous area of Longyan city, Fujian province, a comprehensive promotion technology of traffic safety was proposed through the

investigation of traffic characteristics and traffic accident data and analysis of the traffic flow characteristics, traffic accident characteristics and causes of traffic accidents. The implementation effect was evaluated. The results showed that the application of the comprehensive promotion technology of traffic safety reduced the risk of traffic accidents and improved the traffic safety level.

References

1. John F (2006) Preliminary report for the I-215 corridor redevelopment project are. Amendment No. 1a—lakeview/nuevo sub-area
2. Guan YC (2013) Study on traffic safety design technology of though highway in villages and small towns. Chang'an University
3. Zhao JL (2006) Study on installation of traffic signs on mountain road. Chang'an University
4. Tang CC, He Y (2008) Technology for setting up traffic safety facilities on two-lane highways. China Communications Publishing House
5. Zhang TJ, Tang CC, Kang YX (2012) Based on negative binomial model mountain two-lane highway in rural road safety features. Highw Traffic Technol 29(6):114–119, 136
6. Research Institute of Highway Ministry of Transport, Department of Transportation of Guizhou Province (2015) Technical guide for the implementation of highway safety and life protection engineering (trial implementation). China Communications Press Co., Ltd
7. Ma XD, Shen W, Ma MT, Han FC (2017) Study on an active traffic safety management system for mountain highway in rural area. In: 17th COTA international conference of transportation professionals, pp 4699–4710

Determinants of Long Distance Traveler's Arrival Modes: A Case Study of the Beijing Capital Airport

Zhenhua Mou, Weiwei Liang, Yanyan Chen, Yao Lu and Shaohua Wang

Abstract Before the formal long-distance travel, the travelers usually have four common modes to get to the airport or railway station: metro, airport coach, auto, and taxi. This study was proposed to analyses the determinants of long-distance traveler's arrival modes to the airport. Based on the theory of planned behavior (TPB), a questionnaire survey was designed and conducted to acquire the variable data of the psychological factors that affect urban air passengers' arrival modes. After the pilot survey, Beijing Capital International Airport was chosen to conduct the survey and more than 3700 sample data was acquired. With the sample data, the coefficient relationship between the behavior attitude, subjective norm, perceived behavior control, and behavior intention was analyzed after validating the reliability and validity. The study also employed a structural equation model (SEM) to explore the insight between the determinants and decision. The correlation variables analyses result and path coefficient reveal that the behavior intention of using various travel modes with the subjective norm having the most impact on the behavior intention. The conclusion part explained why the railway was less used than the coach and people love to use taxi or auto to arrive the air terminals. This study reveals the key determinant that influence the choice of behavior.

Z. Mou · Y. Chen (✉) · S. Wang
Beijing Key Laboratory of Traffic Engineering, No.100, Pingleyuan, Beijing, Chaoyang District, China
e-mail: cdyan@bjut.edu.cn

Z. Mou
e-mail: mouzhenhua@sdjzu.edu.cn

S. Wang
e-mail: 79432696@qq.com

Z. Mou · W. Liang · Y. Lu
Shandong JianZhu University, No.1000 Fengming Rd, Jinan, Licheng District, China
e-mail: 2546107559@qq.com

Y. Lu
e-mail: luyao980651@gmail.com

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4_115

Keywords Transport engineering · Theory of planned behavior · Structural equation model · Travel mode · Long distance travel

1 Background and Previous Research

With the improvement of income and living standard, an increasing number of people prefer air travel for long distance trips. According to statistics in 2016, there were 28 civil airports whose throughput capacity was over 10 million passengers [1]. The total throughput of all civil airports in China exceeds 1 billion and grows 11.1% per year. The largest civil airport is Beijing Capital International Airport, which handles more than 94 million passengers. Along with the substantial increase in goods shipped by air, many airports have begun to augment their serving capacity and that hence the landside transportation system suffers greater traffic pressure. The purpose of this study was to optimize the prediction of air passengers' arrival mode, to solve congestion around super large-scale airports, and to support the city's travel demand management strategy.

In the previous studies, researchers studied travel characteristics and laws in different traffic modes mainly from the view of temporal and spatial travel characteristics or traveler socioeconomic characteristics. Chen [2] researched the distribution rate of Beijing New Airport under the compete of high-speed railway and the current airport with thousand of questionnaire data, and the research found four determinants in mode choice decision-making: costs, time, comfort level and flexibility, the conclusion was brought forth based on the logistic regression model. Huang [3] found the air passenger group mobility has a periodicity and is easy to be affected by festival and holidays, and the research was based on an emerging data, the ticket order data. Also, there were so much previous researches that explained the passenger's travel mode decision making from the aggregate model applied in the modal-split phase of the four-step theory [4, 5].

Some researchers analyzed the passenger's activity from the micro perspective in disaggregate model, Lanken, Huneche and Verplanken researched the travel mode choice in different views, such as attitude and habit [6], responsibility and environment [7], context change [8]. The above studies are all exploring the decision-making of travel mode based on micro-individuals, but from different view of consideration. Also, it's the limitation of the existing research, it's necessary to propose a comprehensive model considering subjective and objective and other related factors to evaluate travel mode decision-making behavior. To describe the air passenger's decision process more precisely, an advanced research method must be employed in this study.

Theory of Planned Behavior (TPB) was a method to help people know how his/her decision was made, widely used in the human behavior research, such as consuming [9], healthcare [10] and other related fields since 1990s. The TPB was also employed as an emerging method to explore the insight of human decision on travel mode choice, some researchers keep the original framework unchanged, some extended or

improved the TPB theory framework to adapt to different contents that contain the daily commute, intercity travel [11] and long-distance travel [12]. The air passenger's arrival to airport travel was unique part of the whole travel chain that different from the above travel type, its mode choice was affected by unusual factors such as the travel habits and the accompany person, or even the baggage amount. The TPB theory can just explain this complex decision process.

Starting with the choice of travel modes, the study designed variables in the questionnaire adapt to the TPB framework [13]. First, the social and economic characteristics of survey samples is discussed, subsequently, the reliability and validity of the behavior attitude, subjective norm, perceived behavior control, and behavior intention is tested. Then, descriptive and correlational analyses were conducted on TPB variables data which correspond to the following 4 types of traffic mode: metro, airport coach, auto, and taxi. Finally, it formulated a regression model, Structural Equation Model (SEM), of the behavior intention under various circumstances of the travel behavior attitude, subjective norm, and perceived behavior, which is applicable to apply to TPB in the research on airport travel modes, and the path coefficient between variables and intention was calculated out. In the conclusion and discussion part, the theoretical results and engineering measures that help for improving the passenger's travel experience were put forward.

2 Theoretical Hypothesis and Data Acquire

In TPB there were four important factor terms and one intention term, the behavior attitude (BA) refers to a certain degree to which traveler prefers to execute something, in this study, it refers to the passenger's preference to one of the four modes. The subjective norm (SN) refers to social pressure about pondering over whether a behavior should be implemented, it refers to if the passenger's mode choice was supported by his/her family or friends in this study. The perceived behavior control (PBC) refers to self-assessment or self-perception that an individual performs some specified behavior, it refers to the passenger's success experience and barrier expectation of his/her choice, the more success experience and less barrier expectation he/she has, the value of PBC get bigger. The behavior intention refers to the willingness of an individual to execute a given behavior, in this study it refers to a passenger's selection probability on one of the four modes. Applied in the choice of travel mode, the four partition were measured by one or more items in the form of questionnaire. The behavior attitude was treated as the most important term, it had four measurement items, the subjective norm had two items for the families and friends' attitude on the choice. The perceived behavior control (PBC) is to a certain extent a mode whose arrival time and arrival method are evaluated on account of its convenience, this term was considered to be different with the previous researches, the built environment factor was added into the measurement, the built environment contains population density, development intensity and transit accessibility [14, 15]. In the light of Ajzen's description about the variables in TPB, the following hypotheses are proposed [13]:

(1) The preference of airport passengers for a certain travel mode has a significant positive impact on the willingness to choose.
(2) The social pressure of airport passengers is positively influenced by their willingness to choose a mode.
(3) The convenience of a travel mode has a significant positive impact on the willingness to choose the method among airport passengers.

In this study, the Likert scale was used to design the questionnaire on travelers' behavior intention to get to the airport. It divided the statements about the behavior attitude into two types: positive and negative, and set the grade scores of each sentence according to the tendency of attitude. The answer to the Likert scale of the 4–6 interval was the most common. Berdie considered that a 5-points scale is most reliable in most cases [16, 17]. After the survey, the Likert scale was applied to convert respondents' answers into corresponding marks and then accumulated them as attitude scores. The average scores represented the overall attitude of the evaluation object which means the higher the marks, the more positive the attitude will be.

According to the main variables of TPB, the corresponding measurement items were designed. As shown in Table 1, the 4 modes of travel behavior attitude were measured through 3 specific questions: the comfort index, cost, and time consumption. The subjective norm is represented by the attitude from relatives and friends for a variety of travel modes. The perceived behavior control points to the perception of a specific behavior and a judgement of whether it is easy to execute and is measured by the proficiency of using various travel modes. The behavior intention is measured through the degree of applying willingness among various travel modes.

After pilot survey, the C, D, and E departure zones in Beijing Capital International Airport were chosen as the investigation sites, and over 3700 questionnaires were acquired by random interview, among which, the proportion of passengers on international routes and domestic routes was about 1:3. As the largest airport on a national scale, Beijing Capital International Airport offers divers modes for air passengers. In addition to the four main modes mentioned above, there are tour buses, buses, hotel

Table 1 Questionnaires of TPB

Variables in TPB	Questions
Behavior attitude	The degree of overall satisfaction with certain traffic modes The degree of how comfortable it is in certain traffic modes The degree of perception of the fees in certain traffic modes The feeling of travel time in certain traffic modes
Subjective norm	The degree of support for certain traffic modes from the most important relatives or friends
Perceived behavior control	The degree of hassle in certain traffic modes The degree of easily handling in certain traffic modes during the whole process, including the effect of built environment
Behavior intention	The degree of willingness in using certain traffic modes The degree of willingness in using certain traffic modes under normal conditions

buses, car-hailing service, and other modes. Due to the limit sample, the minority modes were not considered in the research.

3 Statistical Analyses of Social Characteristics and Test of Reliability and Validity

In the valid samples, 57.69% were male and 42.31% were female. The age was divided into 4 stages: 25 and under, 26–35, 36–45, 46–55, and 56 and above, respectively coded from 1 to 5. As can be seen that from Fig. 1, the main composition of passengers were young and middle-aged (26–45) enterprise and institution staffs who accounted for 67%. From the macro statistics, taxis accounting for 48% was the most important travel mode to air passengers, then were car, airport bus and metro. It was found from the survey of the amount of baggage, that most airport passengers carry large bags and over 87% have one or more, and they are faced with the problem of transferring the baggage. The taxi mode relied on flexible door-to-door [18] service was becoming more popular. Since the survey was conducted during the summer vacation, a certain proportion of students were the respondents which also affected the distribution of income to some extent from which an annual income of less than one hundred thousand yuan accounted for the highest proportion, 45%. The proportion of one hundred thousand to two hundred thousand was 39%, and income phases decrease in turn which reflects the fact that overall income level of air passengers is relatively high.

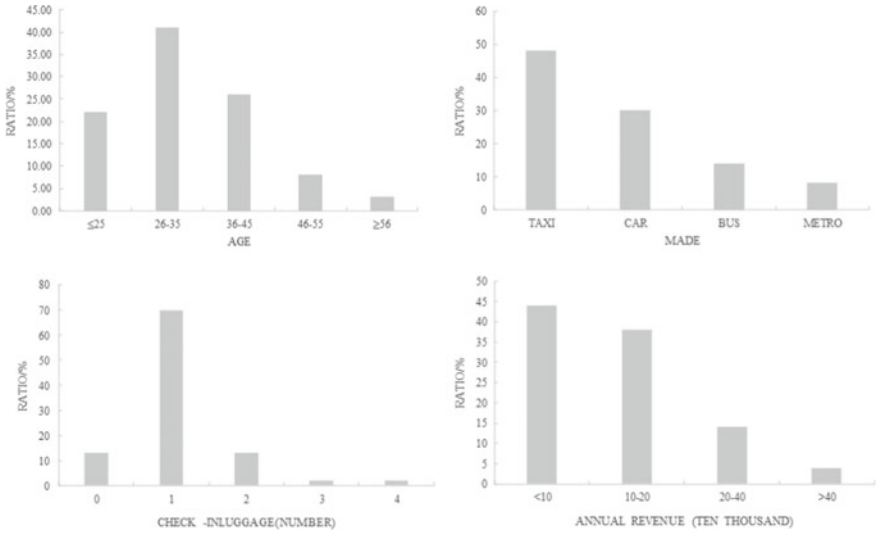


Fig. 1 Social characteristics statistics of the survey samples

Reliability refers to the stability of a measuring tool, which represents the proximity of repeated measurements [13]. The evaluation index is the reliability coefficient. The higher the coefficient, the more consistent, stable, and reliable the result will be. The coefficient of reliability is also called Cronbach's α coefficient, and Xi (i = 1, 2, …, n) is the observation of the problem item. The formula of Cronbach α is as follows:

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum s_i^2}{s_{total}^2}\right], \quad s_{total}^2 = \mathrm{var}\left(\sum x_i\right) \\ s_i^2 = \mathrm{var}(x_i) \tag{1}$$

The average variance extracted (AVE) can be used to describe the convergent validity. AVE reflects how many variances explained by latent variable are derived from the items. When the AVE value is over 0.50, it means the latent variables have good convergent validity. Here only the relevant inspection value of the car mode is given.

It can be seen from Table 2 that in factor analysis, all the principal component factors with eigenvalues greater than 1 are unique, and the most variance contributions reach 60%. Through first-order confirmatory factor analysis, it was found that variables in the model and the corresponding factor loadings of the problem items are both greater than 0.5, and the Z values are over 1, indicating that the sample has good validity.

In order to test the consistency and stability of the measurement results, the variables' Cronbach's α coefficient were calculated [10]. It was found that the index of the taxi mode: the behavior attitude and subjective norm were greater than 0.72, as

Table 2 Cronbach's α and AVE of the variables in TPB

Traffic mode	Variables	Principal component factor analysis		Cronbach's α coefficient	AVE value
		Percentage of total variance explained	Eigenvalue		
Car	Behavior attitude	65	1.23	0.58	0.58
	Subjective norm	71	1.34	0.55	0.69
	Descriptive norm	69	1.35	0.51	0.69
	Perceived behavior control	71	1.85	0.69	0.64
	Behavior habit	63	2.63	0.77	0.56
	Behavior intention	61	1.75	0.64	0.61

well as the descriptive norm of the metro mode and the behavior attitude and subjective norm of the airport bus mode which indicated strong internal consistency while the other variables' s Cronbach's α coefficient were 0.50–0.68.

The above principal component factor analysis proves the validity of the unique factor from the observation variable set. According to George and Mallery, the rules on Cronbach's α coefficient was that above 0.9 was excellent, above 0.8 was better, above 0.7 was acceptable, above 0.6 was basically recognized, above 0.5 was poor, and less than 0.5 was unacceptable [13]. The convergence validity AVE value of each variable was greater than 0.5, so the variable had a better convergent validity.

4 Descriptive Statistics of Variables

The mean and variance of various travel modes of related variables with TPB were given in Table 3. They clearly showed that among a variety of travel modes, whether behavior attitude, subjective norm, perceived behavior control, or behavior intention, taxis obtained the highest score which means that respondents had the most preferences for taxi travel. The social environment, and public opinion also tends to recommend the use of taxis. Due to the lower number of transfers, it was very convenient for travelers to carry more baggage and the perceived behavior control of taxis reaches the highest score of all variables in total. Car mode was a similar experience compared to taxi, so it had a relatively high score. Compared with taxi and car travel, the scores of each variable of the metro travel are low so the travelers had a more distinct preference for taxis and car pick-up.

Based on the gender, age, and income of the respondents, the influence of social statistical characteristics on each variable in TPB was analyzed. According to the analysis of the social statistical characteristics and considering that the young population made up a larger proportion, the study divided the age below 35 years old and over 36 years into young and middle-aged groups, 100 thousand's annual income

Table 3 Descriptive statistics of the variables in TPB

Variables	Airport bus		Taxi		Car		Metro	
	Mean value	Variance	Mean value	Variance	Mean value	Variance	Mean value	Variance
Behavior attitude	3.42	1.08	4.16	0.99	4.07	1.11	3.26	1.16
Subjective norm	3.41	1.05	4.19	0.81	3.82	0.99	3.26	1.10
Perceived behavior control	3.44	2.17	4.21	0.88	4.19	2.00	3.48	1.12
Behavior intention	3.07	1.13	4.08	0.80	3.69	1.05	3.00	1.11

was considered as the threshold of high-income groups. The mean values of variables under various social characteristics were shown in Table 4.

From Table 4, among the sexes, there was no significant difference in the scores of various variables at the 95% confidence level. From the specific numerical point of view, the scores of men for a variety of variables are relatively higher than that of females which demonstrates that males more readily accept different travel modes. For the metro mode, the difference of the behavior intention was most remarkable between males and females, which corresponds to the taxi travel mode in that females marked more scores of the behavior attitude and behavior intention which showed that women have more requirements for travel safety and environmental cleanliness.

The same as for gender, on the whole age was not a significant factor for each variable. But it was important to note that in the car mode, the score of the travel behavior intention and subjective norm marked by the middle aged was higher than the youth for 0.29 and 0.24 which is the biggest difference between all the variables in the scoring, which means that middle-aged people tend to travel more by car than youths and they consider that their relatives and friends are more supportive than youths for car traveling.

Although the difference in income level for each variable score was not significant, from the difference of various ways of travel behavior intention, the higher income respondents were more inclined to use taxi and car travel. Especially for the car and airport bus mode, the travel intention score of those having higher incomes was higher than the lower income for 0.16 and 0.18. The behavior attitude includes the degree of comfort, speed, and price factors. The difference between the index of the generalized cost was more sensitive to income while the metro was not.

5 Correlation Analyses of Variables

In order to study the relationship between different trip modes and behavior intention, the Pearson correlation coefficients of each behavior intention scores were calculated, and the null hypothesis was proved by using the t statistic. As shown in Table 5, ** indicates a coefficient of significant correlation at the 0.01 significance level. As can be seen from Table 5, the choice of taxi is significantly related to the behavior intention of the airport bus and car modes, and the taxi and airport bus's correlation coefficient is −0.20. The coefficient is 0.21 with the car mode which means that the stronger the intention of the taxi mode is, the willingness of using the airport bus will be reduced. However, the intention of using taxi and car travel modes has a positive correlation that the competitive relationship does not exist between the behavior intention of taxi and car travel choice. The correlation coefficient between the behavior intention of choosing the metro and the airport bus travel is 0.32, which shows that there is a strong positive correlation between the choice of the two travel modes.

For each travel mode shown in Table 6, the behavior attitude, subjective norm, perceived behavior control, and behavior intention were also analyzed. As can be

Table 4 Analysis of the variables of TPB with respect to sex, age, and salary

Traffic mode	Variables	Sex		Age		Salary	
		Male	Female	Youth	Above middle age	Low	High
Airport bus	Behavior attitude	3.48	3.29	3.43	3.35	3.48	3.30
	Subjective norm	3.44	3.31	3.42	3.35	3.36	3.40
	Perceived behavior control	3.50	3.30	3.46	3.35	3.41	3.40
	Behavior intention	3.16	2.88	3.08	2.98	3.05	3.02
Taxi	Behavior attitude	4.02	4.25	4.14	4.12	4.18	4.08
	Subjective norm	4.17	4.12	4.05	4.27	4.17	4.13
	Perceived behavior control	4.21	4.11	4.16	4.17	4.21	4.16
	Behavior intention	4.04	4.07	3.97	4.13	4.04	4.07
Car	Behavior attitude	4.05	4.02	4.06	4.01	4.08	4.00
	Subjective norm	3.70	3.89	3.68	3.92	3.74	3.84
	Perceived behavior control	4.15	4.15	4.18	4.11	4.18	4.12
	Behavior intention	3.60	3.71	3.52	3.81	3.57	3.73
Metro	Behavior attitude	3.27	3.17	3.20	3.24	3.26	3.18
	Subjective norm	3.33	3.10	3.21	3.23	3.23	3.21
	Perceived behavior control	3.50	3.38	3.41	3.48	3.46	3.42
	Behavior intention	3.11	2.81	2.90	3.06	3.00	2.95
P values tested by independent sample t		0.0835		0.2008		0.2891	

Table 5 Correlation analysis between intentions of different travel modes

Travel modes	Airport bus		Taxi		Car	
	Coefficient	*P* value	Coefficient	*P* value	Coefficient	*P* value
Taxi	−0.20**	0.00				
Car	−0.10	0.05	0.21**	0.00		
Metro	0.32**	0.00	−0.04	0.47	−0.02	0.61

Note: ** means $p < 0.01$

Table 6 Correlation analysis between variables of different travel modes

Travel modes	Variables	Behavior attitude		Subjective norm		Perceived behavior control	
		Coefficient	*P* value	Coefficient	*P* value	Coefficient	*P* value
Airport bus	Subjective norm	0.28	0.00				
	Perceived behavior control	0.27	0.00	0.48	0.00		
	Behavior intention	0.40	0.00	0.48	0.00	0.52	0.00
Taxi	Subjective norm	0.40	0.00				
	Perceived behavior control	0.33	0.00	0.59	0.00		
	Behavior intention	0.35	0.00	0.57	0.00	0.45	0.00
Car	Subjective norm	0.40	0.00				
	Perceived behavior control	0.42	0.00	0.38	0.00		
	Behavior intention	0.39	0.00	0.71	0.00	0.29	0.00
Metro	Subjective norm	0.55	0.00				
	Perceived behavior control	0.53	0.00	0.61	0.00		
	Behavior intention	0.55	0.00	0.65	0.00	0.56	0.00

seen from Table 6, all variables are significantly related. Among the users of the airport bus, the subjective norm and perceived behavior control is strongly related to the behavior intention. In the taxi and metro mode, the subjective norm, perceived behavior control, and behavior intention has a strong relationship. In the car mode, the subjective norm and behavior intention directly show a strong correlation. Generally, the subjective norm and perceived behavior control have more influence on the behavior intention in different ways of travel [16, 18].

6 Structural Equation Model Applied on Decision Making

In order to study the relationship between various psychological variables of arrival at the airport within the framework of TPB, a structural equation model was constructed and a model estimation was carried out, Structural Equation Modeling (SEM) is a very good method in social science research, it can deal with the multiple causes and results, or encounter unobservable variables (latent variables), which are problems that cannot be solved well by traditional statistical methods. In this study, the SN, Attitude and PCB were treated as the exogenous variables that only affect other variables, but not affected by other variables. The Intention was treated as the endogenous variable that only affected by the exogenous variables.

The model is defined as shown in Fig. 2. Among them, SN is the subjective norm, Attitude is the behavior attitude, PBC is the perceived behavior control, and Intention is the behavior intention. Assuming I is an observation index vector, and η is a psychological latent variable vector

$$l_y = \Lambda_y \eta + \varepsilon \tag{2}$$

$$l_x = \Lambda_x \xi + \delta \tag{3}$$

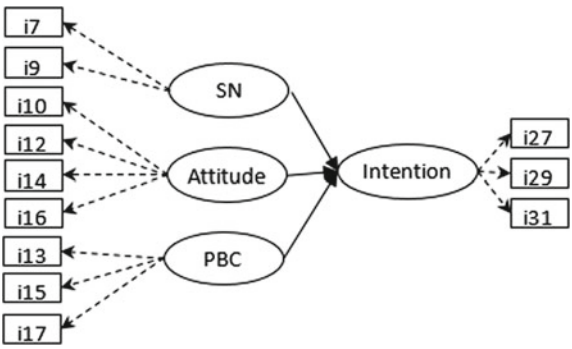


Fig. 2 Detailed path analysis of the original TPB

Table 7 Structural equation model's goodness-of-fit

Travel modes	χ^2/df	RMSEA	CFI	TLI
Airport bus	1.825	0.053	0.972	0.954
Taxi	2.210	0.075	0.963	0.938
Car	1.872	0.049	0.978	0.965
Metro	2.573	0.061	0.977	0.963

In these equations η is an endogenous latent variable which represents Intention, ξ is an exogenous latent variable vector which represents SN, Attitude, and PBC, I is an observable variable, and ε and δ are errors.

The structural equation

$$\eta = B + \Gamma\xi + \zeta \tag{4}$$

B represents the endogenous effect between variables, Γ represents the effects of exogenous latent variables on endogenous latent variables, and ζ represents endogenous latent variables that can not be explained by the error term.

The study selected degrees of freedom of chi square (x^2/df), root mean squared error of approximation (RMSEA), comparative fit index (CFI), and the Tucker-Lewis index(TLI) as the fitting evaluation indexes between the research on the structural equation and survey data.

First, the behavior attitude, subjective norm, and perceived behavior control were chosen as exogenous latent variables based on the travel mode and behavior intention as endogenous latent variables. Based on the survey data, the structural equation model of the 4 modes: airport bus, taxi, car, and metro, was established. The model fitting evaluation index is shown in Table 7.

It can be seen from Table 7 that the value of χ^2/df of various travel modes' structural equation model is within 1.825–2.573, showing that the hypothesis model and the fitting degree between the sample data can be accepted, and RMSEA, CFI, and TLI are within the acceptable range of the model fitting. In general, the fitting degree of the structural equation model in the study is within the acceptable range and can be further analyzed by using the results.

The estimation results of the structural equation model for various travel modes are shown in Table 8. For different travel modes, the behavior attitude, subjective norm, and perceived behavior control have different impacts on the behavior intention. For the airport bus mode, hypotheses 1 and 2 have both been validated. For the taxi and car modes, hypothesis 2 has been validated. For the metro mode, all the hypotheses are proved. Specifically, for the airport bus mode, the behavior attitude and subjective norm have a significant positive effect on the behavior intention for their normalized path coefficients are 0.23 and 0.53, and the behavior attitude, subjective norm and perceived behavior control are not significantly correlated. Also, the explanation of the perceived behavior control for the behavior intention is not correlated. This shows that the preference of the airport bus and the perception of

Table 8 Estimation of the model

Travel mode	Itinerary	Normalized path coefficient	Z value
Airport bus	Behavior attitude → Behavior intention	0.23	2.44*
	Subjective norm → Behavior intention	0.53	6.37**
	Perceived behavior control → Behavior intention	0.15	1.43
Taxi	Behavior attitude → Behavior intention	0.06	0.54
	Subjective norm → Behavior intention	1.09	10.63**
	Perceived behavior control → Behavior intention	0.02	1.80
Car	Behavior attitude → Behavior intention	0.06	0.59
	Subjective norm → Behavior intention	0.90	13.70**
	Perceived behavior control → Behavior intention	0.01	0.04
Metro	Behavior attitude → Behavior intention	0.17	2.27*
	Subjective norm → Behavior intention	0.29	2.72**
	Perceived behavior control → Behavior intention	0.50	4.52**

Note: * means $P < 0.05$, ** means $p < 0.01$

the subjective norm significantly influence the selection of the airport bus and the subjective norm embodies the expectation of the travel mode choice behavior of traveler's friends. From the standardized path coefficient, the subjective norm shows the greatest social explanatory power for the choice behavior of the airport bus. For the taxi and car modes, only the subjective norm has a significant positive influence on the behavior intention and their path coefficients are 1.09 and 0.90 respectively, and the behavior attitude, subjective norm, and perceived behavior control are positively related. For the metro mode, the behavior attitude, subjective norm, and perceived behavior control have a significant positive effect on the behavior intention, and the 3 exogenous latent variables are significantly positively correlated. The perceived behavior control of the behavior intention of the standardized path coefficient is 0.50, the maximum, and the subjective norm and behavior attitude are 0.28 and 0.17 respectively, indicating that those using the metro mode are more ready to ignore the impact of inconvenient factors (carrying more baggage) on the journey. In general, the subjective norm variable has a significant positive effect on the behavior intention of the travel mode choice which indicates that travelers will consider more

social expectations or pressure when choosing travel modes. Travel preferences and the perception of how convenient the mode will be influence the behavior intention which is also consistent with Ajzen's interpretation of TPB that in the field of behavior research, the influence of exogenous variables of TPB on the behavior intention or even behavior is different [17].

7 Conclusion and Further Discussion

TPB and SEM were introduced into the study of arrival mode on air terminal and the effects of various variables on the behavior of arriving at the airport were systematically investigated. The research showed that the behavior attitude, subjective norm, perceived behavior control, and behavior intention had significant differences between different travel modes, with the difference between the choice of the behavior attitude of the taxi and airport bus being the most significantly different. Not all demographic characteristics had remarkable effects on various variables, and each feature for one or several travel modes has a certain sensitivity.

The scale of the designed travel mode choice of the behavior attitude, subjective norm, perceived behavior control, and behavior intention had good reliability and validity which is suitable for the study of the airport travel mode choice behavior. The relationship paradigm of each TPB variable was successfully applied to the travel mode choice behavior, and the subjective norm in a variety of modes had the most notable impact on the choice of the behavior intention.

The results revealed that there was different variable play important role in different modes, some descriptive characteristics can be explained by the results. In the metro mode, the perceived behavior control also showed a significant impact. The study used the original plan behavior of 3 basic variables in TPB, the behavior attitude, subjective norm, and perceived behavior control, and new variables such as built environment, past behavior, behavior habit, descriptive norm, and expected regret [19] were also introduced into other researches in the field of traffic behavior. It revealed that the metro has less split because the transfer obstacle the travelers who have more luggage, and the auto and taxi can provide the door-to-door service to avoid the negative effect. To raise the split of metro mode, there should make some improvement on the transfer channel to fit for the passengers who have much baggage. The variables designed with good reliability and validity of the scale in the air passenger's arrival mode need to be screened one by one. Future research is aimed at further analyzing the explanation and forecast of new variables about the mode choice of the behavior intention and actual behavior based on the survey data.

There are also some shortcomings in the research, the study just put the sight on traveler themselves subjectively but ignore the objective circumstance's effects, such as the built environment factor which has a growing literature in the relative researching. The relative research has a bright future but a long way to go.

Acknowledgements This study was supported by the Humanities and Social Science Funds of the Ministry of Education (Grant:19YJC630124).

References

1. Van de Vijver E, Derudder B, Witlox F (2016) Air passenger transport and regional development: cause and effect in Europe. Promet-Traffic Transp 28(2):143–154
2. Chen W, Huang J, Yan M et al (2012) Research on the characteristics of Beijing passengers and the distribution rate of new airport. Prog Geogr 31(10):1360–1368
3. Huang F, Peng J, Yan M et al (2016) Analyses of characteristics of air passenger group mobility behaviors. Acta Physica Sinica 65(22):228901
4. Scheiner J (2010) Interrelations between travel mode choice and trip distance: trends in Germany 1976–2002. J Transp Geogr 18(1):75–84
5. Limtanakool N, Dijst M, Schwanen T (2006) The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium-and longer-distance trips. J Transp Geogr 14(5):327–341
6. Lanken B, Aarts H, Van Knippenberg A et al (1994) Attitude versus general habit: antecedents of travel mode choice. J Appl Soc Psychol 24(4):285–300
7. Hunecke M, Blöbaum A, Matthies E et al (2001) Responsibility and environment: ecological norm orientation and external factors in the domain of travel mode choice behavior. Environ Behav 33(6):830–852
8. Verplanken B, Holland RW (2002) Motivated decision making: effects of activation and self-centrality of values on choices and behavior. J Pers Soc Psychol 82(3):434
9. Madden TJ, Ellen PS, Ajzen I (1992) A comparison of the theory of planned behavior and the theory of reasoned action. Pers Soc Psychol Bull 18(1):3–9
10. Godin G, Kok G (1996) The theory of planned behavior: a review of its applications to health-related behaviors. Am J Health Promot 11(2):87–98
11. Jing P, Juan ZC, Zha QF (2014) Incorporating psychological latent variables into travel mode choice model. China J Highw Transp 27(11):85–92
12. Xin L (2014) Research on the passenger travel choice of moderate/long-distance based on the planned behavior theory. MD thesis, Southwest Jiaotong University, Chengdu
13. Ajzen I (2011) The theory of planned behavior: reactions and reflections. Psychol Health 26(9):1113–1127
14. Ma C, Hao W, Wang A et al (2018) Developing a coordinated signal control system for urban ring road under the vehicle-infrastructure connected environment. IEEE Access 6:52471–52478
15. Sun X, Wandelt S, Hansen M et al (2017) Multiple airport regions based on inter-airport temporal distances. Transp Res Part E 101:84–98
16. Merkert R, Beck M (2017) Value of travel time savings and willingness to pay for regional aviation. Transp Res Part A: Policy Pract 96:29–42
17. Shevlin M, Miles JNV, Davies MNO (2000) Coefficient alpha: a useful indicator of reliability? Pers Individ Differ 28(2):229–237
18. Ma C, He R, Zhang W (2018) Path optimization of taxi carpooling. PLoS One 13(8):e0203221
19. Rhodes RE, Plotnikoff RC, Spence JC (2004) Creating parsimony at the expense of precision? conceptual and applied issues of aggregating belief-based constructs in physical activity research. Health Educ Res 19(4):392–405

Author Index

W. Wang et al. (eds.), *Green, Smart and Connected Transportation Systems*,
Lecture Notes in Electrical Engineering 617,
https://doi.org/10.1007/978-981-15-0644-4