

Wuhong Wang
Yanyan Chen
Zhengbing He
Xiaobei Jiang *Editors*

Green Connected Automated Transportation and Safety

Proceedings of the 11th International
Conference on Green Intelligent
Transportation Systems and Safety

Lecture Notes in Electrical Engineering

Volume 775

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany

Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Stanford University, Stanford, CA, USA

Yong Li, Hunan University, Changsha, Hunan, China

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institut of Telecommunications, Universität Stuttgart, Stuttgart, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Walter Zamboni, DIEM - Università degli studi di Salerno, Fisciano, Salerno, Italy

Junjie James Zhang, Charlotte, NC, USA

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, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (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)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at <https://link.springer.com/bookseries/7818>

Wuhong Wang · Yanyan Chen · Zhengbing He ·
Xiaobei Jiang
Editors

Green Connected Automated Transportation and Safety

Proceedings of the 11th International
Conference on Green Intelligent
Transportation Systems and Safety

 Springer

Editors

Wuhong Wang
Department of Transportation Engineering
Beijing Institute of Technology
Beijing, China

Zhengbing He
Beijing University of Technology
Beijing, China

Yanyan Chen
Beijing University of Technology
Beijing, China

Xiaobei Jiang
Beijing Institute of Technology
Beijing, China

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-16-5428-2

ISBN 978-981-16-5429-9 (eBook)

<https://doi.org/10.1007/978-981-16-5429-9>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Modeling and Analysis of the Impact of Charging Electric Charged Vehicles on the Grid Load in the Expressways	1
Dazhou Li and Wei Gao	
Research on the Time Interval of Discrete Traffic Flow Model	21
Liang Rui, Shan Zheng, and An Zeping	
An Approach to Analyze Commuters' Transfer Characteristics on the Public Transportation Network in an Incomplete Data Environment	37
Shichao Sun	
Understanding Metro-to-Bus Transfers in Nanjing, China Using Smart Card Data	51
Miaoyan Zhang, Dongwei Liu, Yanjie Ji, and Yang Liu	
Research and Optimization of Vehicle Braking Efficiency Under Different Working Conditions Based on AMESim	69
Yiyi Wei, Baoliang Ji, Ziyuan Xiong, Jingqiang Liang, Wei Liu, and Zhen Huang	
Effects of the Layout of Exhaust Port on the Performance of Smoke Control for Ultra-Wide Immersed Tube Tunnel	83
Ying Li, Shenglin Mu, Xiaoyang Ni, Zhongchen Niu, and Kaixuan Tang	
Research on Safety Management of Inland River Dangerous Chemicals Transportation System Based on System Dynamics and Principal Component Analysis	97
Yu Zhao and Zhiyong Lv	
Mechanical Properties of Concrete Under the Alternate Carbonation and Freezing–Thawing Actions	109
Jin Ran, Qi Feng, Jinxi Zhang, Lin Wang, and Honglin Cao	

Research on the Top-Level Design of Urban Road Intelligent Promotion	121
Jin-dong Zhou	
Predicting Red Light Running Violation Using Machine Learning Classifiers	137
Muhammad Zahid, Arshad Jamal, Yangzhou Chen, Tufail Ahmed, and Muhammad Ijaz	
Bi-Objective Subway Timetable Optimization Considering Changing Train Quality Based on Passenger Flow Data	149
Hanlei Wang, Peng Wu, Yuan Yao, and Xingxuan Zhuo	
Wind Tunnel Tests on Aerodynamic Noise from the Head Car of a High-speed Train	163
Hai Deng, Yigang Wang, Qiliang Li, Zhe Shen, Yang Gao, and Jin Lisheng	
Invulnerability Simulation of Urban Agglomeration Traffic Network Cascading Failure Under Random Attack Strategy	177
Chengbing Li, Zhicheng Yang, and Zhenyu Liu	
Impact of Intelligent Networking on Vehicles Exiting at Urban Intersections	195
Jun Liu and Shu-Bin Li	
Safety Evaluation of Expressway Linear Operation Based on Design Consistency	219
Ke Chen, Jian-Feng Jiang, Jia-Jia Tian, Yun-You Liu, and Yanli Ma	
Forecast Model of Comprehensive Transport Freight Index Considering Macro Economy in China	231
Yanli Ma, Jieyu Zhu, and Yining Lou	
An Optimization Model of Passenger Transportation Structure Based on Land Use and Energy Consumption	245
Fenghe Yu, Yining Lou, Lifei Han, and Jieyu Zhu	
A Review of Motor Vehicle Driver Fatigue Research—Based on L3 Automatic Driving System	257
Hanying Guo and Yuhao Zhang	
Analysis of Compliance Between Cape Town Agreement and China’s Fishing Vessel Standards	271
Majing Lan, Chaoyu Ruan, and Muqiu Zhou	
Analysis of Non-motor Vehicle Traffic Flows and Traffic Conflicts on Isolated Bicycle Lane	275
Lin Wang, Jin Ran, Zhonghua Wei, Shi Qiu, and Shaofan Wang	

Parameter Matching and Simulation Analysis on Power System of Wheel-Legged Vehicle 291
Xizheng Chen, Xiaobei Jiang, Lijin Han, Wenlin Yu, Jiawen Guo, and Wuhong Wang

An Online Impedance Algorithm for All-Climate Batteries Used on Electric Vehicle 301
Tao Zhang, Xiaoxia Sun, Jun Tian, and Ning Zhao

A Research of Eco-driving Strategies Based on CVIS Considering Lane Change Behaviors 309
Jingyu Zhang, Chengcheng Xu, and Pan Liu

Research on Charging Station Planning Technology Based on Pure Electric Taxi SOC Driving Range in Northern Cold Area City 321
Haoran Liu, Bing Wang, Xuejing He, Junyou Sun, and Kuo Zhao

Model Construction and Simulation Analysis of Vehicle Collision on Steep Road Section in Mountainous Area 333
Xuezhao Fang, Bing Wang, Haoran Liu, and Kuo Zhao

Research on the Influence of Freeway Truck Mixing Rate on Traffic State 349
Fengchun Han, Na Li, Dan Zhao, and Yumeng Zhang

Experimental Method to Establish Vehicles' Acceleration/Deceleration Speed Parameter Thresholds for Urban Roads in Beijing 363
Chenjing Zhou, Xinyue Nie, Zihe Lin, Jian Rong, and Jianming Ma

A Comparative Study on Institutions and Mechanisms for Investigating and Handling Maritime Accidents in China and Major Developed Countries 377
Minglu Ma and Lili Jiang

The Status Quo of China's Implementation of International Maritime Conventions and Suggestions and Countermeasures for Its Participation in Global Governance 389
Minglu Ma and Guobo Wang

Multi-source Data Congestion Recognition Study Based on Fuzzy Pattern Recognition 397
Junzhuo Li, Wenyong Li, and Bin Guo

Research on Classifiers Used to Identify Dangerous Goods Transportation Vehicles 411
Haodong Zhang, Qian Cheng, Kuikui Feng, Xiaobei Jiang, and Wuhong Wang

The Police Perspective on Driving Styles in China: An Interview Study	423
Siyuan Zhou, Xu Sun, Bingjian Liu, and Gary Burnett	
Simulation Methods of Ultrasonic Guided Wave Propagation Mechanism and Sensor Deployment in Rail	437
Miao Zhang, Guoqiang Cai, and Ye Zhang	
Passing Ability Impact Factors Analysis of Heavy Tractor	447
Hao Zhang, Cheng-qiang Zong, Chuan-jin Ou, and Yue Li	
Roller Opposite Forces Type Automobile Brake Tester Key Technical Parameters' Experimental Research	457
Yuanpeng Liu, Xiaoping Tong, and Hao Zhang	
Trends Analysis of Domestic Waterway Traffic Safety Production Based on Regional Characteristics in China	479
Guobo Wang, Guanquan Chu, and Chao Han	
Study on Route Optimization of Seasonal Hot Spots	491
Haixia Zhao, Ting Tong, and Jiangping You	
Research on Energy-Saving Driving for Transport Vehicles Considering Actual Load	509
Chunming Li, Tao Zhang, Xiaoxia Sun, Chunming Shao, and Guozhu Wang	
Research on the Coupling Relationship Between Passengers and Vehicles Based on Anylogic Simulation	523
Jiahui Li, Siyuan Hao, Chengwu Jiao, Nale Zhao, and Keman Wu	
Coevolution Dynamics Model of Urban Multimodal Transportation System Under Limited Exploitable Land Resources	533
Rong Li, Weili Kong, Liwen Zhang, and Peng Shu	
Traffic Congestion Improvement Scheme of the Corridor Between Beijing Deputy Center and Central Urban Area	543
Chuanjiao Sun and Panyi Wei	
Research on the Plan of Container Train Operation in Sea-Rail Intermodal Transportation Under the Uncertain Demand	549
Zhiyi Lei, Zhikuan Ma, Jéssica Saturnino Sabino de Sousa, and Yifan Zhu	
Research on Setting of Warning Signs for Highway Tunnels in High Altitude Areas	563
Jian-you Zhao, Wei-guo Xie, Wan-li Han, and Wan-jiang Guo	
Analysis on the Road Capacity: Connected and Automated Vehicle Platoons in the Mixed Traffic Flow	577
Ke Ma and Hao Wang	

Real-time Trajectory Prediction of Unmanned Aircraft Vehicles Based on Gated Recurrent Unit 585
 Rong Tang, Zhao Yang, Jiahuan Lu, Hao Liu, and Honghai Zhang

The Capacity Model of Urban Multimode Transportation Network Based on Combined Travel Mode 597
 Guozhu Cheng and Chuanmiao Zhou

Autonomous Train Dispatch Method for Regional Rail Transit Based on Deep Reinforcement Learning 607
 Wei Wei, Ling Liu, Jun Liu, and Bo Zhang

Research on Energy Delivery of a Series–Parallel Hybrid Electric Vehicle Under Different Driving Conditions 621
 Chunming Li, Xiaoxia Sun, Chunming Shao, Lining Yang, Chenglong Shu, Danhua Niu, and Guozhu Wang

Information Extracting and Analysis on Safe Operation of Online Car-hailing Based on Text Mining: The Case of China 635
 Si Yang Liu, Jian Rong, Chen Jing Zhou, Yi Wang, and Zi He Lin

The Method of Division and Classification Quantification for Snow and Ice Removal on Airport Pavement 649
 Ya-ping Zhang, Yu Zhang, Xiao-qing Xing, Ye Chen, Nan Yang, and Jian Mao

Research on Object Detection and Location Using Semantic Segmentation 665
 Huijie Zhu, Yan Cai, Yingchun Guo, Xiang Feng, and Wenya Zhang

General Standard Model for Safety Risk Assessment of Highway Bridge Construction 675
 Yuqian Wang and Dianliang Xiao

Modeling Simulation and Stability Evaluation Analysis of Double-Trailer Combination 689
 Yue Li, Hao Zhang, Hongwei Zhang, Chengqiang Zong, Chuanjin Ou, and Xueli Zhang

An Improved Comprehensive Evaluation Method of Road Passenger Transportation Safety 707
 Guoliang Dong, Fujia Liu, Haiying Xia, and Chaozhou Chen

Vehicle Detection at Night Based on the Feature of Taillight and License Plate 723
 Guosheng Ma, Mante Cai, Guanliang Chen, and Zhixiao Li

Spatial Integration Based on Big Data—A Case Study on the Integration of Nanjing and Ma’anshan 745
 Xiaofeng Lou and Hong Zhang

A Method of Household Car Ownership Prediction Using Ordered Probit Model 759
Guangzheng Yao, Yanyan Chen, Kaijun Cui, Donghui Xu, and Jiarui Liu

Long Short-Term Memory Neural Network for Traffic Speed Prediction of Urban Expressways Using Floating Car Data 773
Deqi Chen, Xuedong Yan, Shurong Li, Xiaobing Liu, and Liwei Wang

Design and Implementation of Automobile Test Monitoring Management System Based on PDA Terminal 789
Hong Jia, Hao Li, and Hai-ying Xia

A Three-layer Optimal Distribution Problem for Electric Vehicle Charging Stations 797
Di Chen, Xinyu Yu, Linghan Li, and Yuyang Zhou

Analysis of Mixed Vehicle Traffic Flow at Signalized Intersections Based on the Mixed Traffic Agent Model of Autonomous-Manual Driving Connected Vehicles 817
You Ren, Shan Jiang, Guan Yan, Hongmei Shan, Huiying Lin, Zhilong Zhang, Liangzhe Wang, Xuesheng Zheng, and Jiaqi Song

Driving Velocity Tracking Error Analysis of Different Broadcast Methods Under Green Light Optimal Speed Advisory System 831
Shanshan Guo, Tao Zhang, and Yongsheng Liu

Mechanical Modeling and Analysis Based on Fine Drill 845
Kuikui Feng, Haodong Zhang, Qian Cheng, and Wuhong Wang

Author Index 855

Modeling and Analysis of the Impact of Charging Electric Charged Vehicles on the Grid Load in the Expressways



Dazhou Li and Wei Gao

Abstract With the development of the energy grid, the number of electric charging cars will gradually increase on the highway. Highways must not only provide enough natural gas and fuel for vehicles, but also have enough electricity to provide a large number of electric charging cars in the future. When electric charging cars on the highway become the main body of the vehicles, the grid load problem caused by the charging of electric charging cars on the highway will be a problem that must be faced and resolved. The present study is aimed to analyze and model the power load on the highways, combined with electric vehicles penetration scenarios, in order to produce geospatial information for potential energy demand with charging stations in the highway. According to the fluid theory, the efficient and timely prediction of the grid load generated by the charging of electric charging vehicles has been proposed. It will not only help the construction of the grid along the highway, but also contribute to the stability, speed and smoothness of highway traffic. This paper selects Shenda Expressway in China as a realistic research object, and analyzes and models the power load demand of charging stations from 3 aspects: peak-to-valley difference, load balancing and voltage excursion.

Keywords Expressways · Electric vehicles · Charging management · Power load · Peak-to-valley difference · Load balancing · Voltage excursion

1 Introduction

Transportation electrification offers two important advantages. First, it reduces local carbon emissions and fossil-fuel dependence [1]. Second, it shifts energy needs toward a power system that is increasingly able to leverage energy produced from renewable sources [2, 3]. Electric vehicles, in particular, allow for emission reduction in urban areas and, due to their use patterns in urban environments, can potentially

D. Li · W. Gao (✉)

College of Computer Science and Technology, Shenyang University of Chemical Technology, Shenyang 110016, China

e-mail: gaowei_syuct@foxmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,

Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_1

operate as flexible electric loads to support the operation of power systems and the integration of renewable energy [4–6]. Therefore, research on electric vehicles and their related technologies has become the focus of attention of governments, automakers, and energy companies [7].

With the popularity of electric vehicles, the rapid charging of a large number of electric vehicles disorderly connected to the grid can lead to a sharp increase in the load on the user side of the grid. It will cause a great negative impact on the stable of the grid operation and result in a decrease in energy load. Furthermore, as the main concentration of electric vehicles in the future, the highway is also bound to generate a large amount of electrical load. The electrical load will not only cause a large impact on the power grid, but also need to renovate the existing power grid along the highway to ensure smooth highway traffic.

Therefore, how to achieve reasonable and orderly charging management of a large number of electric vehicles has become a hot topic of research. How to analyze and model the grid load generated by electric vehicles on highways will be an issue that the transportation and power sectors need to consider together in the future. In order to solve above problem, this paper conducts joint modeling of the traffic flow of electric vehicles on the expressway, in which the charging service provided by the charging station, and analyzes the calculation of the electrical load required for electric vehicles to travel in concentrated traffic based on the classical fluid theory and queue service theory.

2 Related Works

Research effort on electric vehicles' charging in highway has largely focused on three main issues:

- (1) Optimal placement of charging stations: Napoli et al. [8] proposed an optimal locations of electric vehicle infrastructures in a highway network. Zhang et al. [9] proposed a highway charging station planning method based on the dynamic traffic flow simulation. Wang et al. [10] investigated the siting and sizing problem of fast charging stations in a highway network, where the budget constraint and the service capacities of charging spots are considered. Covic et al. [11] detailed a new topology for the track pad, consisting of two largely coplanar, partially overlapping coils positioned such that there was no mutual inductance between them. Li and Xie [12] developed an optimal charging station location model and method for locating charging stations for electric vehicles in intercity highway networks with the aim of minimizing the detour cost of the driving population. Jinglin et al. [13] calculated and identified the optimal positions of the charging stations for electric vehicles in the Italian highway network. Zhao and Burke [14] studied the feasibility of the deployment of renewable hydrogen fueling at Highway Safety Roadside Rest Areas

and the integration of the stations with the electricity grid to lower the infrastructure cost and to accelerate the usage of renewable energy in the California transportation sector.

- (2) Development of electric vehicles' charging scheduling algorithm: In this part, the appropriate scheduling algorithm and a process by which individual electric vehicles make correct decisions about their charging locations are both research targets. In [15], Simulation of the number of electric vehicles is carried out applying the distribution of residential load, and site planning for charging station is put into operation to maximize the operating income of charging station. Victor [16] proposed the Vehicle-Originating-Signals approach for charging control of a fleet of electric vehicles in an electricity distribution network. Gusrialdi et al. [17] addressed the system-level scheduling problem and the individual control problem, while requiring only distributed information about electric vehicles and their charging at service stations along a highway. They also developed a strategy to coordinate the queues among the charging stations, with only local information about traffic flows and the status of electric vehicle charging stations along a bidirectional highway, so that excessively long waiting times had be avoided [18]. Reuter et al. [19] provided a comparison of two applied approaches—one coverage oriented approach and an optimization based method—which can be easily transferred to other regions due to comparably low data requirement. Razo and Jacobsen [20] proposed a smart scheduling approach for electric vehicles to plan charging stops on a highway with limited charging infrastructure to minimize the total travel time for each electric vehicle.
- (3) Analyzing and modeling the load of the grid along the expressway: This paper falls into this aspect. Some of these problems have already been considered in the literature. In [21], the probabilistic model of charging load is studied. Xie et al. [22] built a model of calculating the charging demands of electric vehicles running in highway to satisfy the travelling demands of large scale of electric vehicle in intercity highway. The literatures [7, 23–25] analyzed the impact of the large number of electric vehicles accessing the grid and suggested that load management strategies should fully consider the distribution of electric vehicle charging loads in time and space in order to avoid overloading of the power system during peak hours. Based on the charging demand and load characteristics of electric vehicle, modeling and calculation method is proposed in [26, 27].

3 Modeling Method

3.1 Modeling of Electric Vehicle Flow in Expressways

There are multiple lanes on the freeway. Vehicles moving in two directions have different forward speeds. Assume that a section of highway is represented by a

length of horizontal line. The left end is the starting point of the highway and has 0 distance in the coordinate system. The distance of the vehicle traveling to the right from the left starting point is represented as x . The speed of the vehicle traveling to the right is denoted by $v(x, t)$ at time t in the distance x from the left origin.

In general, $v(x, t) \geq 0$, $x \geq 0$, $+\infty \geq t \geq 0$. The analysis and calculation on the reverse lane are consistent with mentioned above, but the directions are opposite. The right end point is the starting point of the leftward movement in the highway. In later content, we take the vehicle moving from left to right as an example. The analysis and modeling methods of a vehicle moving from right to left are completely equivalent in the same way.

In this paper, the methods, based on the theory of fluid flow in [28–30], are used to estimate the number of charging cars in each charging station on the highway. The following assumptions are made: the amount of power of an electric vehicle that has been fully charged before coming into a highway is the upper limit of the battery capacity. It allows electric vehicles to run one-way journey on the highway. If a fully-charged electric vehicle cannot run the entire highway, it must stop and charge in the halfway. The assumption is rationality that will not only avoid burden the driver, but also resist the congestion on the highway.

We define the following 4 variables. At time t , $A(x, t)$ is the number of electric cars that need to be charged on the highway $(0, x)$ section. In the range of time $(0, t)$, $B(x, t)$ is the number of electric vehicles that need to be charged out of the range of $(0, x)$, but still on the other section of the expressway. In the range of time $(0, t)$, $C(x, t)$ is the number of electric vehicles that need to be charged coming into the highway $(0, x)$ section from any entrance of the $(0, x)$ section. Within the time range $(0, t)$, $D(x, t)$ is the number of electric vehicles that will charge the power and leave the highway in any exit in the $(0, x)$ segment. $D(x, t)$ consists of two types of vehicles. One type of vehicles arrive at the destination and leave from the highway in any exit in the $(0, x)$ segment. The other type of vehicles enter the charging station for charging due to insufficient power and return to the highway to the destination after charging power.

In the $(0, x)$ section of the highway and the time range $(0, t)$, the net flux of the electric vehicle to be charged is equal to the sum of $A(x, t)$ and $B(x, t)$. Furthermore, it also equals to the difference between $C(x, t)$ and $D(x, t)$, as shown in Formula (1).

$$A(x, t) + B(x, t) = C(x, t) - D(x, t) \quad (1)$$

Assuming $A(x, t)$, $B(x, t)$, $C(x, t)$, and $D(x, t)$ are continuous in time and space in Formula (1), the following conclusions can be drawn.

First, at time t on the freeway point x , the density $a(x, t)$ of the vehicle to be charged is equal to the partial derivative of $A(x, t)$ for distance x .

$$a(x, t) = \frac{\partial A(x, t)}{\partial x} \quad (2)$$

Secondly, at time t and freeway point x , the flow $b(x, t)$ of the vehicle to be charged is equal to the partial derivative of $B(x, t)$ for time t .

$$b(x, t) = \frac{\partial B(x, t)}{\partial t} \quad (3)$$

According to Formula (1–3), we can get Formula (4)

$$\frac{\partial a(x, t)}{\partial t} + \frac{\partial b(x, t)}{\partial x} = \frac{\partial^2 C(x, t)}{\partial t \partial x} - \frac{\partial^2 D(x, t)}{\partial t \partial x} \quad (4)$$

According to the hydromechanics theory in [31], the flow at the outflow end of a pipe is equal to the fluid density multiplied by the flow velocity of the fluid at the outflow end of the pipe. When we consider the section of highway $(0, x)$ as a section of pipeline, the left end of the highway, in the coordinate system 0, is the entrance of the pipeline. The point x of the highway is the exit of the pipeline. The vehicle fluid flow $b(x, t)$ at the pipe exit can be regarded as the flow of the vehicle that needs to be charged at the x point of the highway. The density $a(x, t)$ of the vehicle flow fluid at the outlet of the pipeline can be regarded as vehicle fluid density that needs to be charged at the x point of the highway. The velocity of the flow fluid $v(x, t)$ can be regarded as the speed of vehicle that needs to be charged. Therefore, we obtain the Formula (5) by [30, 32].

$$b(x, t) = a(x, t) \cdot v(x, t) \quad (5)$$

Substitute Formula (5) into Formula (4).

$$\begin{aligned} & \frac{\partial a(x, t)}{\partial t} + \frac{\partial (a(x, t) \cdot v(x, t))}{\partial x} \\ &= \frac{\partial^2 C(x, t)}{\partial t \partial x} - \frac{\partial^2 D(x, t)}{\partial t \partial x} \end{aligned} \quad (6)$$

Formula (6) expands.

$$\begin{aligned} & \frac{\partial a(x, t)}{\partial t} + \frac{\partial a(x, t)}{\partial x} \cdot v(x, t) + \frac{\partial v(x, t)}{\partial x} \cdot a(x, t) \\ &= \frac{\partial^2 C(x, t)}{\partial t \partial x} - \frac{\partial^2 D(x, t)}{\partial t \partial x} \end{aligned} \quad (7)$$

Because speed can be considered as the product of displacement and the inverse of time, $v(x, t)$ can be written as

$$v(x, t) = \frac{\partial x(t)}{\partial t} \quad (8)$$

According to the partial derivative calculation rule [30], the Formula (9) is obtained

$$\frac{da(x, t)}{dt} = \frac{\partial a(x, t)}{\partial t} + \frac{\partial a(x, t)}{\partial x} \frac{dx(t)}{dt} \quad (9)$$

Formula (9) and Formula (8) are substituted into Formula (7) to obtain Formula (10)

$$\begin{aligned} & \frac{da(x, t)}{dt} + \frac{\partial v(x, t)}{\partial x} \cdot a(x, t) \\ &= \frac{\partial^2 C(x, t)}{\partial t \partial x} - \frac{\partial^2 D(x, t)}{\partial t \partial x} \end{aligned} \quad (10)$$

When the time is in the range of $(0, t)$, there are $C(x, t)$ electric vehicles that need to be charged in the Section $(0, x)$ of the highway from any entrance in the Section $(0, x)$. If there are n highway entrances in the $(0, x)$ segment, each entrance records $c(k)$ that is the number of electric vehicles entering the $(0, x)$ segment at every k seconds. The second-order partial derivatives of $C(x, t)$ for time t and position x can be obtained according to [30] as illustrated in Formula (11).

$$\frac{\partial^2 C(x, t)}{\partial t \partial x} = \sum_{i=1}^n c_i(k) \quad (11)$$

$D(x, t)$ is composed of two types of vehicles. One type of electric vehicles reach the destination and apart from any one of the exits in the segment $(0, x)$. The number is represented by $E(x, t)$. The other type of electric vehicles enter the charging stations due to insufficient power, which will return to the highway to continue to its destination after fully charging. The number of latter type of electric vehicles is represented by $F(x, t)$.

$$D(x, t) = E(x, t) + F(x, t) \quad (12)$$

If there are m highway exits in the $(0, x)$ segment, each exit records $e(k)$ that is the number of electric vehicles leaving the highway at every k seconds. The second-order partial derivatives of $E(x, t)$ versus time t and position x can be obtained according to Formula (13) by [30].

$$\frac{\partial^2 E(x, t)}{\partial t \partial x} = \sum_{i=1}^m e_i(k) \quad (13)$$

If there is a charging station in the $(0, x)$ section, which is built in a rest area at the location x , the maximum number of electric vehicles parking in the charging station is denoted by Max . At time t , $h(x, t)$, $h(x, t) \leq Max$, is the number of electric vehicle in the charging station. If $h(x, t)$ equals to Max , the charging station is saturated for parking. Only one electric vehicle fully has been charged and left, the next one can be allowed to enter for charging. When the number of vehicles are fully charged at

time t is presented by $p(x, t)$, we have Formula (14).

$$\frac{\partial^2 F(x, t)}{\partial t \partial x} = p(x, t) \quad (14)$$

When $h(x, t) < Max$ and $h(x, t) + a(x, t) \leq Max$, the number of vehicles that need to be charged at position x is denoted by $a(x, t)$.

$$\frac{\partial^2 F(x, t)}{\partial t \partial x} = a(x, t) \quad (15)$$

When $h(x, t) < Max$ and $h(x, t) + a(x, t) > Max$, the number of vehicles that the charging station at position x at time t is given as follow.

$$\frac{\partial^2 F(x, t)}{\partial t \partial x} = Max - h(x, t) + p(x, t) \quad (16)$$

Substituting Formula (11–16) into Formula (10), we get

$$\left\{ \begin{array}{l} \frac{da(x, t)}{dt} + \frac{\partial v(x, t)}{\partial x} \cdot a(x, t) \\ = \sum_{i=1}^n c_i(k) - \sum_{i=1}^m e_i(k) - p(x, t) \\ \text{when } h(x, t) = Max \\ \frac{da(x, t)}{dt} + \frac{\partial v(x, t)}{\partial x} \cdot a(x, t) \\ = \sum_{i=1}^n c_i(k) - \sum_{i=1}^m e_i(k) - a(x, t) \\ \text{when } \begin{cases} h(x, t) < Max \\ h(x, t) + a(x, t) < Max \end{cases} \\ \frac{da(x, t)}{dt} + \frac{\partial v(x, t)}{\partial x} \cdot a(x, t) \\ = \sum_{i=1}^n c_i(k) - \sum_{i=1}^m e_i(k) - Max + h(x, t) - p(x, t) \\ \text{when } \begin{cases} h(x, t) < Max \\ h(x, t) + a(x, t) < Max \end{cases} \end{array} \right. \quad (17)$$

In Formula (17), except for $a(x, t)$, the rest can be obtained from actual highway information. The partial derivative of speed $v(x, t)$ for the displacement x can be obtained according to the restrictions on vehicle speed in the traffic rules of different sections of the highway. For example: the maximum speed of different lanes is from 110 to 90 km/h, and the minimum speed is limited to 80–60 km/h. Moreover, it also can be considered that the vehicle speed is a random value at the displacement x and time t between the maximum and the minimum speed limit of the road sections. The

calculation of partial derivative can be operated by using the difference of velocity at adjacent positions. Therefore, under the conditions of given x and t through the numerical solution of the partial differential equation, the value of $a(x, t)$ can be obtained by Formula (17).

For solving the partial differential equation and setting the boundary conditions as mentioned above, we adopt $a(0, t)$ and $v(0, t)$ to represent the initial density and speed of the electric vehicle that needs to be charged at the left end of the expressway. The variable q_0 is the arrival rate of vehicles passing through the left end of the expressway.

$$a(0, t) = v(0, t) \cdot q_0 \quad (18)$$

After calculating $a(x, t)$ through Formula (17), the number of electric vehicles that need to be charged at a given time t and position x can be obtained in Formula (19).

$$\left\{ \begin{array}{l} \frac{\partial^2 F(x, t)}{\partial t \partial x} = p(x, t), \\ h(x, t) = Max \\ \frac{\partial^2 F(x, t)}{\partial t \partial x} = a(x, t), \\ \left\{ \begin{array}{l} h(x, t) < Max \\ h(x, t) + a(x, t) \end{array} \right. \\ \frac{\partial^2 F(x, t)}{\partial t \partial x} = Max - h(x, t) + p(x, t), \\ \left\{ \begin{array}{l} h(x, t) < Max \\ h(x, t) + a(x, t) \end{array} \right. \end{array} \right. \quad (19)$$

3.2 Modeling of Charging Station Queues in Expressways

There are multiple lanes on the freeway, with vehicles moving in two directions, each with a different forward speed. In this paper, the M/M/s/K queue theory [32] can be used to estimate the number of electric vehicles that need to be charged in a charging station located at x at a given moment t .

The electric cars that need to be recharged arrive in Poisson flow with an average of λ cars per minute, i.e., the time that the electric cars that need to be recharged arrive at the charging station successively obeys the negative exponential distribution of the parameter λ .

$$\lambda = \frac{\partial^2 F(x, t)}{\partial t \partial x} \quad (20)$$

The charging station can only park up to Max electric cars waiting to be charged, and the electric cars waiting to be charged are lined up in a queue, the longest queue being Max , using the first-come first-served principle. When the electric vehicle arrives, if that charging station is full, the electric vehicle must be driven to the next charging station.

$$\lambda_i = \begin{cases} \lambda = \frac{\partial^2 F(x,t)}{\partial t \partial x}, & i = 0, 1, 2, \dots, Max \\ 0, & i > Max \end{cases} \quad (21)$$

The charging station is provided with s charging piles. The charging time of the electric vehicle obeys a negative exponential distribution with the parameter $\mu = p(x, t)$.

The charging station has s charging piles. The electric vehicle charging time distribution is a negative exponential distribution of μ .

$$\mu_i = \begin{cases} i \cdot \mu = i \cdot p(x, t), & 0 \leq i < s \\ s \cdot \mu = s \cdot p(x, t), & s \leq i \leq Max \end{cases} \quad (22)$$

According to [33], we used W to represent the charging power of a charging pile and γ to represent the battery capacity of an electric vehicle that needs to be supplemented. Generally, the value of γ can reach 0 to 4.88 kWh. According to Formula (22), $p(x, t)$ can be calculated using given W , s and γ .

According to [33], W denotes the charging power of 1 charging stake and γ denotes the battery capacity that needs to be replenished for electric vehicles. Generally, γ values can reach between 0 and 4.88 kWh. According to Formula (23), $p(x, t)$ can be computed using the given W , s and γ .

$$\frac{p(x, t)}{s} = \frac{\gamma}{W} \quad (23)$$

In order to make the queue system stable, the following inequality should be satisfied, according to the M/M/s/K queue theory:

$$\frac{\lambda}{s \cdot \mu} < 1 \quad (24)$$

When time t is given, the minimum value of the number of charging piles in the charging station located x on the highway is

$$s > \frac{1}{p(x, t)} \frac{\partial^2 F(x, t)}{\partial t \partial x} \quad (25)$$

When s meets the conditions, the queue system is in a stable state. The probability that there are i electric cars in the queue waiting to be charged is represented by p_i .

$$p_i = \begin{cases} \frac{p_0}{i!} \left(\frac{\lambda}{\mu}\right)^i, & 0 \leq i \leq s \\ \frac{p_0}{s! \cdot s^{i-s}} \left(\frac{\lambda}{\mu}\right)^i, & s \leq i \leq Max \end{cases} \quad (26)$$

For convenience, ρ and ρ_s are defined respectively.

$$\begin{cases} \rho = \frac{\lambda}{\mu} = \frac{1}{p(x,t)} \frac{\partial^2 F(x,t)}{\partial t \partial x} \\ \rho_s = \frac{\lambda}{s \cdot \mu} = \frac{1}{s \cdot p(x,t)} \frac{\partial^2 F(x,t)}{\partial t \partial x} \end{cases} \quad (27)$$

For calculating p_i , the value of the initial state p_0 is a hyper-parameter.

$$p_0 = \begin{cases} \left(\sum_{i=0}^{s-1} \frac{\rho^i}{i!} + \frac{(\rho)^s \cdot (1 - (\rho_s)^{Max-s+1}}{s!(1-\rho_s)} \right)^{-1}, & \rho_s \neq 1 \\ \left(\sum_{i=0}^{s-1} \frac{\rho^i}{i!} + \frac{(\rho)^s \cdot (Max-s+1)}{s!} \right)^{-1}, & \rho_s = 1 \end{cases} \quad (28)$$

Considering that the charging station space has a maximum Max , the average number of occupied charging piles (which is also the average number of charging electric vehicles receiving charging service) is N_s , which can be calculated according to the full probability formula.

$$\begin{aligned} N_s &= \sum_{i=0}^{s-1} i \cdot p_i + \sum_{i=s}^{Max} s \cdot p_i \\ &= p_0 \left(\sum_{i=0}^{s-1} \frac{i \cdot (\rho)^i}{i!} + \sum_{i=s}^{Max} \frac{(\rho)^i}{s!(s)^{i-s}} \right) \\ &= p_0 \rho \left(\sum_{i=0}^{s-1} \frac{(\rho)^i}{i!} + \sum_{i=s}^{Max} \frac{(\rho)^i}{s!(s)^{i-s}} - \frac{(\rho)^{Max}}{s!(s)^{Max-s}} \right) \\ &= \rho \left(1 - \frac{(\rho)^{Max}}{s!(s)^{Max-s}} p_0 \right) \\ &= \rho(1 - p_k) \end{aligned} \quad (29)$$

So far, at a given time t , the average power consumption W_p of the grid load in the charging station located x on the highway can be calculated by the product of N_s and W .

$$W_p = N_s \cdot W \quad (30)$$

4 Experiment and Analysis

4.1 Dataset

We take the Shenda Expressway as an example. When the electric charging car becomes the main type of vehicles traveling on the Shenda Expressway, the electrical loads required by the charging stations located in the service area of the Shenda Expressway can be calculated through the relevant data analysis. The calculation can provide a scientific and theoretical basis for the future grid construction and traffic dispatching of the Shenda Expressway (Fig. 1).

The Shenda Expressway is the first eight-lane expressway in China. It is 348 km long. When it was completed in 1990, it had four lanes, with full interchange traffic. The widening and transformation started in 2002 and was completed in 2004. It is an eight-lane expressway with a subgrade width of 42 m. The design speed is 120 km/h. Shenda Expressway has a total of 348 km. The total length of Shenda Expressway

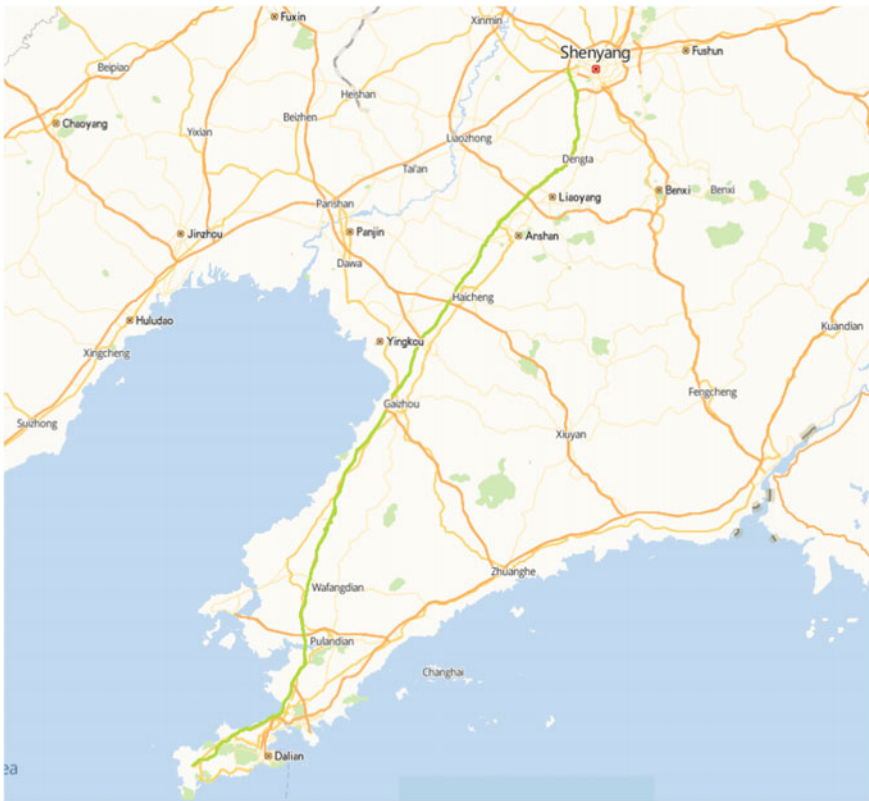


Fig. 1 The Shenda expressway in northeast China is marked with a green line on the map

before the renovation was 375 km. After the renovation, it started from Campbell Interoperability in Shenyang in the north and ended in Houyan Village in Dalian in the south. Shenda Expressway has a total of 38 entrances and exits and 6 service areas, which list in Table 1.

4.2 Experimental Environment and Parameter Settings

There are six service areas, “Jingquan Service Area”, “Ganquan Service Area”, “Xihai Service Area”, “Xiongyue Service Area”, “Fuzhou River Service Area” and “Sanshilibao Service Area” in the Shenda Expressway. Each of the six service areas has a charging station. The six service areas divide the Shenda Expressway into five sections: Shenyang Campbell Interoperability-Jingquan Service Area, Jingquan Service Area-Ganquan Service Area, Ganquan Service Area-Xihai Service Area, Xihai Service Area-Xiongyue Service Area, Xiongyue service area-Fuzhou service Area, Fuzhou river service area-Sanshilibao Service Area. Within each segment, there are multiple exits and entrances. However, charging stations are only set up in the above six rest areas.

Taking the Shenyang Campbell Interoperability-Jingquan Service Area as an example, a horizontal line is used to represent the Shenyang Campbell Interoperability-Jingquan Service Area. Campbell Interoperability is the left starting point of the highway moving to the right, with coordinate 0. The length of the vehicle traveling to the right from the left starting point is represented by x , the speed of the vehicle traveling to the right is represented as $v(x, t)$, and the length of the vehicle traveling to the right is x , at time t from the left starting point of Campbell Interoperability.

For the other example, to calculate the grid load of charging stations located in Jingquan Service Area through the proposed method in Sect. 3, the following parameters will be set up as follow. Assume $v(x, t)$ obeys the distribution ranges from 60 to 120 km/h with an average value of 90 km/h. Normally, $v(x, t)$ can be reset according to traffic rules, such as: bends and gates. The pit stop area is 30 km/h. There are $n = 4$ highway entrances in the section of Campbell Interoperability-Jingquan Service Area. Each entrance records the number of vehicles entering the highway, $c(k)$, in every $k = 60$ s. There are $m = 4$ highway exits. Each exit records the number of vehicles leaving the highway, $e(k)$, in every $k = 60$ s. In the Campbell Interoperability, the density of electric vehicles is $a(0, t)$ and speed is $v(0, t)$. The vehicle arrival rate q_0 can be estimated according to [32].

The charging station located in Jingquan Service Area adopts the following parameter settings. The charging station can accommodate a maximum of 100 electric vehicles and is equipped with 50 charging piles. The maximum charging power of each charging pile is $W = 70$ kW. Generally, charging electric vehicles need to supplement the battery capacity γ , which obeys a normal distribution of 0–4.88 kWh, with the average value 2.5 kWh. The remaining 5 sections in the Shenda Expressway use the same setting method. The only difference is that the average value of γ will

Table 1 The name of Shenda expressway entrance and service area

No.	Entrance and service area	Distance from Shenyang (km)	No.	Entrance and service area	Distance from Shenyang (km)
1	Campbell interoperability	0	23	Gaizhou entrance	177
2	Sujiatun entrance	5	24	S19 Zhuanggai expressway	182
3	Sujiatun South entrance	9	25	Shagangzi entrance	185
4	Shilihe entrance	20	26	Sturgeon entrance	195
5	Lighthouse entrance	30	27	Xiongyue service area	204
6	Jingquan service area	36	28	Liguan entrance	216
7	North entrance of Liaoyang	48	29	Jutun entrance	242
8	Liaoyang entrance	53	30	Fuzhou River service area	257
9	G91 Liaoning central ring road	59	31	Wafangdian entrance	261
10	Entrance to Liaoyang county	65	32	Laohutun entrance	271
11	Anshan entrance	77	33	Wafangdian South entrance	282
12	South Anshan entrance	88	34	Turret entrance	292
13	Ganquan service area	97	35	S12 leather long high speed	297
14	Nantai entrance	103	36	Bay North entrance	303
15	Haicheng entrance	112	37	Shihe entrance	309
16	Xiliu entrance	118	38	Entrance to thirty miles	316
17	G16 Danxi expressway	121	39	Sanshibibao service area	320
18	Huzhuang entrance	134	40	S23 Dayao Bay Shugang expressway	323
19	S21 fuying expressway	144	41	Jiuli service area (Shenyang direction)	331
20	Yingkou entrance	147	42	Golden state entrance	339
21	Yingkou South entrance	169	43	G11 crane big G15 Shenhai expressway	342
22	Xihai service area	172	44	Dalian entrance	348

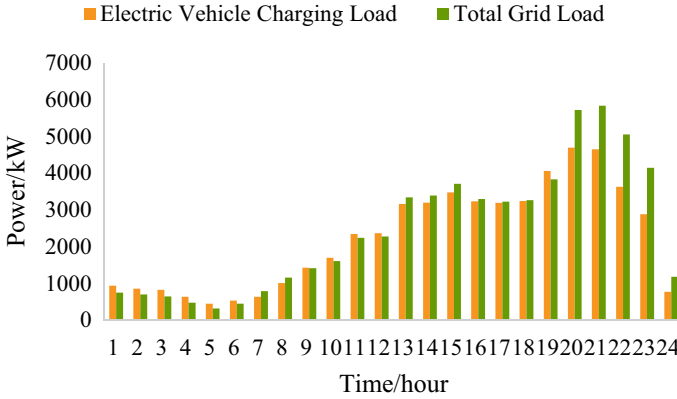


Fig. 2 Electric vehicle charging load and total grid load during a day in No. 30, Fuzhou River service area in the Shenda expressway

gradually increase as the section far away from Campbell Interoperability, because the probability of charging is increasing.

4.3 Result Analysis

Monte Carlo methods can be used to solve any problem with the given probabilistic interpretation. Usually, it tends to follow a particular pattern. Firstly, define a range of possible inputs. The possible inputs contains the constants and the parameters, which have been analyzed in Sect. 3.1. Secondly, inputs will be randomly selected from a probability distribution in the range, illustrated in Sect. 3.2. Thirdly, the inputs will be performed a deterministic distribution and computation. Lastly, aggregate the results. The simulation results are shown in Figs. 2, 3, 4, 5 and 6.

4.3.1 Effect of Electric Vehicle Charging on the Peak-To-Valley Difference of Power Grid

Figure 2 shows the charging load data of electric vehicles and the grid load data that is the total load data of the system after being integrated into the grid in No. 30, Fuzhou River Service Area for one day. When the electric vehicle charging load is calculated as shown in Fig. 2, it is added to the original load of the power grid to calculate the total system load after the electric vehicle is disorderly charging. Furthermore, Fig. 3 indicates two picks of charging load in the No. 39 Sanshilibao Service Area during a day. In Fig. 3, the peak hours of the charging load of the power grid are mainly concentrated at 12:00–16:00 and 19:00–21:00, especially during the period of 19:00–21:00.

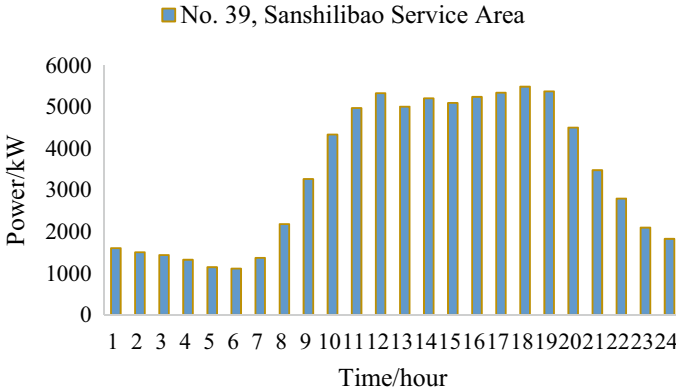


Fig. 3 Electric energy load distribution of No. 39, Sanshilibao service area in the Shenda expressway during a day

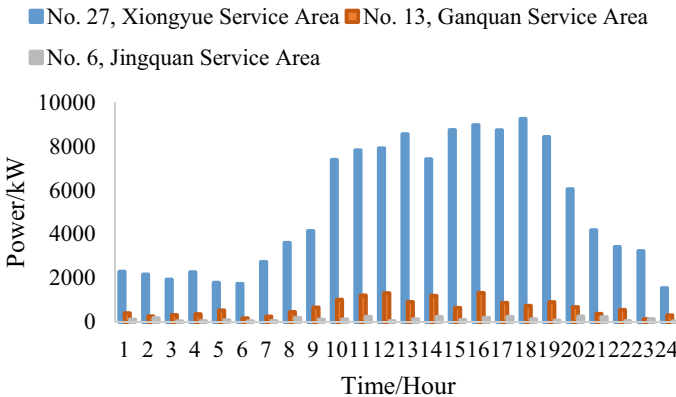


Fig. 4 Electric energy load distribution of No. 6, Jingquan service area, No. 13, Ganquan service area, and No. 27, Xiongyue service area in the Shenda expressway during a day

Figures 2 and 3 demonstrate that the effect of electric vehicle charging on peak-to-valley difference of power grid need be dealt with. Under no guidance charging, the electric vehicle is allowed to charge at any time. The charging load is mainly concentrated during the peak period of the original system load. Because the electricity price is also very high at this moment, the user charging at this time not only will fail to achieve the effect of “peak cutting”, but also bear higher charging costs. The most effective way should be applied to order the charging time of electric vehicles.

In order to measure the impact of charging load on the peak-to-valley difference ratio of the power grid, we perform the calculation and obtain a result that indicates as follow. When the electric vehicle is not connected, the original load peak-to-valley difference is 41.73%. After the electric vehicle is disorderly connected to charge, the

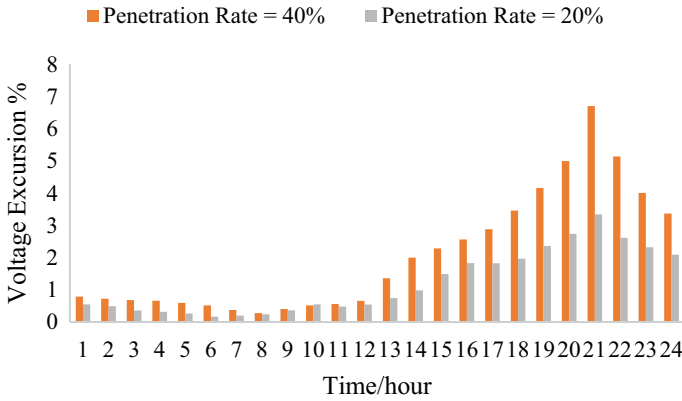


Fig. 5 Voltage excursion during a day in No. 30, Fuzhou River service area in the Shenda expressway

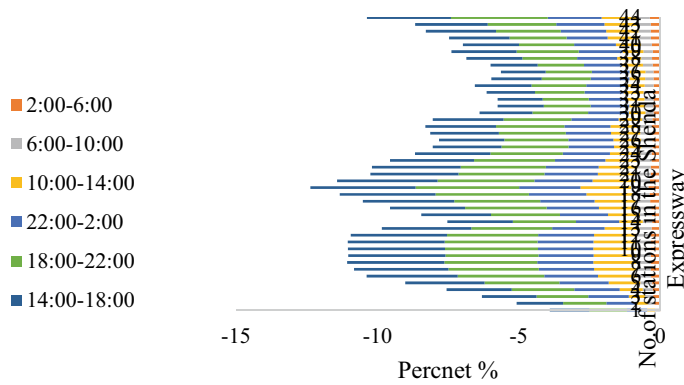


Fig. 6 Variation on voltage of service area and entrance in the Shenda expressway during a day into 6 slots

load peak-to-valley difference becomes 46.51%, which has an increase of 4.78% than no charging. It means if electric vehicles connected to the grid for charging in a disordered state, it will not only enlarge the peak-to-valley difference of the load and exacerbate the fluctuation of the load curve, but also will increase the peak-to-valley difference ratio of the grid load. Moreover, disorder charging also introduces a threat to safe parking. Therefore, it is necessary to guide the charging behavior of electric vehicles, which can not only reduce the impact on the grid load, but also use it as a special energy storage.

4.3.2 Effect of Electric Vehicle Charging on the Load Balancing of Power Grid

As shown in Fig. 4, the huge variation is observed in the profile of charging stations, which indicates that the load balancing of charging stations should be improved. No. 27, Xiongyue Service Area needs more power supply than No. 6, Jingquan Service Area and No. 13, Ganquan Service Area for a day. If we want to ensure the grid has a more balanced line load rate, the high-generation installed nodes and high-load nodes must have a stronger connection with the grid in No. 27, Xiongyue Service Area.

To handle the effect of electric vehicle charging load balancing between the different service areas, the line transmission power margin should be increased and the transmission of power should be ensured. The relatively balanced current distribution at a small economic cost must be considered before charging station setting up. It not only guarantees the technical economy of the scheme, but also rationalizes the line capacity margin ratio. Above all else, it improves the safety of the grid in various operating modes.

4.3.3 Effect of Electric Vehicle Charging on the Voltage Excursion of Power Grid

Excessive voltage excursions cause a series of adverse effects on the grid, which includes the residents' life hazard, shorten the service life of power equipment, and negative effects on the grid itself. The voltage excursion of the power grid in 24 h is shown in Fig. 5. It analyzes the impact of electric vehicle connection on the grid voltage excursion in No. 30, Fuzhou River Service Area. The penetration rate 20% and 40% are connected to the electric vehicle charging load, respectively.

With reference to Fig. 2, the charging load in Fig. 2 remains basically unchanged during the period of 0:00–9:00, remains relative stable in the range of 9:00–13:00, and reaches its peak at 21:00. Similarly, in Fig. 5 the voltage excursion remains unchanged from 0:00 to 9:00, and the voltage excursion remains relative stable from 9:00 to 13:00. This far, even though the charging load of electric vehicles and the impact on the grid voltage have gradually increased, it is still under control.

Between 14:00 and 16:00, the charging load and the voltage excursion increased significantly. At 21:00, the voltage excursion reaches its peak. When the penetration rate reaches 40%, the voltage excursion increases to 10.5%. During this period, not only the peak-to-valley difference ratio for a day is greater, but also the load of the power grid and the impact of the electric vehicle on the voltage of the power grid are both the largest.

Therefore, adjusting the charging time of the electric vehicle is an effective way to improve its impact on the voltage excursion of the power grid. If the charging time is changed from 18:00–23:00 to 1:00–8:00, the voltage excursion will be small, which will affect the power grid less.

Figure 6 shows the voltage variation through the day of the service areas and entrances listed in Table 1, which contains 44 stations in the Shenda Expressway. If not been ordered, the charging time of electric vehicle is relatively concentrated, which will impact a lot on the distribution power grid on voltage loss, harmonic and the balance of three phase. We consider voltage loss as the mainly impact. The loads is added into the grid and calculated according the flow theory. We classified the statistic results during a day, and divide them into 6 slots: 2:00–6:00, 6:00–10:00, 10:00–14:00, 14:00–18:00, 18:00–20:00, and 20:00–2:00 for calculating the voltage loss separately. The largest deviation is observed during 14:00–18:00.

As a result, to avoid the harm of voltage excursion on the grid, methods should be taken to hold the voltage. Many vehicle to grid methods have been developed recently, which treat the battery of electric vehicles as containers to hold the frequency. Distributed power, such as wind and solar energy, in the 44 stations listed in Table 1, is also helpful of clipping the picks. Moreover, the degree of voltage excursion varies from charging station should be considered in different areas.

5 Conclusions

This paper proposes a methodology to analyze impacts of electric vehicle using the example of public available information for the Shenda Expressway in China. The study demonstrates the importance of electric vehicle mobility for the entire energy demand. Our results indicate potential benefits to Shenda Expressway and increase power independence. The classic fluid theory and queue theory are used to analyze and model the electricity load demand along the Shenda Expressway after the charging station is built in the service area along the Shenda Expressway.

According to the method provided in this article, the electric energy load of the service area along the Shenda Expressway while providing electric vehicle charging services, and the electricity required by the entire Shenda Expressway can be analyzed. While the final goal would be to analyze impacts to the distribution and grid, more details such as traffic information, service areas supplied by electric pipes, arrival rate, and traffic rules would be required to model results and the overall impact to the power system.

In the future, the construction of the power grid along the Shenda Expressway will be combined with the construction of the energy grid. The transport artery and the energy artery will gradually merge with each other. The research in this article will provide a certain theoretical basis for the common development of the two.

Acknowledgements We would like to take this opportunity to acknowledge the hard work of all our editors and also the efforts of all the anonymous reviewers who have provided very valuable and helpful comments. We thank them all for their invaluable contributions, and for helping to our research.

Patents This section is not mandatory, but may be added if there are patents resulting from the work reported in this manuscript.

Author Contributions Conceptualization, methodology, formal analysis, data curation, project administration, writing—original draft preparation and funding acquisition, Dazhou Li; editing, validation, resources, software and writing—review, visualization and supervision, Wei Gao.

Funding Project supported by The Science and Technology Funds from Liaoning Education Department (No. LQ2017008); Doctoral Research Startup Fund Project of Liaoning Province (No. 2016011968); Liaoning Higher Education Society 13th Five-Year Plan Shenyang University of Chemical Technology General Topics (GHYB160163).

Conflicts of Interest The authors declare no conflict of interest.

References

1. Boko M, Niang I, Nyong A, Vogel A, Githeko A, Medany M, Yanda PZ (2018) Africa climate change 2007: impacts, adaptation and vulnerability: contribution of working group II to the fourth assessment report of the intergovernmental panel on climate change
2. Bilgin B, Magne P, Malysz P, Yang Y, Pantelic V, Preindl M, Emadi A (2015) Making the case for electrified transportation. *IEEE Trans Transp Electrific* 1(1):4–17
3. Emadi A (2011) Transportation 2.0. *IEEE Power Energy Mag* 9(4):18–29
4. Sovacool BK, Hirsh RF (2009) Beyond batteries: an examination of the benefits and barriers to plug-in hybrid electric vehicles (PHEVs) and a vehicle-to-grid (V2G) transition. *Energy Policy* 37(3):1095–1103
5. Han S, Han S, Sezaki K (2010) Development of an optimal vehicle-to-grid aggregator for frequency regulation. *IEEE Trans Smart Grid* 1(1):65–72
6. Saber AY, Venayagamoorthy GK (2010) Intelligent unit commitment with vehicle-to-grid—a cost-emission optimization. *J Power Source* 195(3):898–911
7. Song Y, Yang X, Lu Z (2010) Integration of plug-in hybrid and electric vehicles: experience from China. In: *IEEE PES general meeting*, pp 1–6. IEEE
8. Napoli G, Polimeni A, Micari S, Dispenza G, Antonucci V (2019) Optimal allocation of electric vehicle charging stations in a highway network: part 2 the Italian case study. *J Energy Storage* 26:101015
9. Zhang J, Zhou H, Liu S, Liu C, Zeng S, Jiang L, Liu H (2018) Highway charging station plan based on dynamic traffic flow simulation. *IOP Conf Ser: Earth Environ Sci* 170(3):032014
10. Wang Y, Shi J, Wang R, Liu Z, Wang L (2018) Siting and sizing of fast charging stations in highway network with budget constraint. *Appl Energy* 228:1255–1271
11. Covic GA, Kissin ML, Kacprzak D, Clausen N, Hao H (2011) A bipolar primary pad topology for EV stationary charging and highway power by inductive coupling. In: *2011 IEEE energy conversion congress and exposition*, pp 1832–1838. IEEE
12. Li J, Xie C Optimally locating charging stations for electric vehicles in intercity highway networks
13. Jinglin H, Yunpeng L, Jun Y, Yang Z, Tianhua H, Shengtao F, Liang C (2018) Planning of electric vehicle charging station on highway considering existing service areas and dynamic traffic simulations. In: *2018 China international conference on electricity distribution (CICED)*, pp 2645–2649. IEEE
14. Zhao H, Burke A (2016) Deployment of sustainable fueling/charging systems at California highway safety roadside rest areas
15. Du A, Hu Z, Song Y, Wu J (2011) Distribution network planning considering layout optimization of electric vehicle charging stations. *Power Syst Technol* 35(11):35–42
16. Victor J (2016) Coordinated electric vehicle charging in residential and highway environments (Doctoral Dissertation, Technische Universität München)

17. Gusrialdi A, Qu Z, Simaan MA (2014) Scheduling and cooperative control of electric vehicles' charging at highway service stations. In: 53rd IEEE conference on decision and control, pp. 6465–6471. IEEE
18. Gusrialdi A, Qu Z, Simaan MA (2017) Distributed scheduling and cooperative control for charging of electric vehicles at highway service stations. *IEEE Trans Intell Transp Syst* 18(10):2713–2727
19. Reuter-Oppermann M, Funke S, Jochem P, Graf F (2017) How many fast charging stations do we need along the German highway network. In: Proceedings of the EVS30 international battery, hybrid and fuel cell electric vehicle symposium, pp 9–11. Stuttgart, Germany
20. del Razo V, Jacobsen HA (2016) Smart charging schedules for highway travel with electric vehicles. *IEEE Trans Transp Electric* 2(2):160–173
21. Chen G, Mao Z, Li J, Wang D, Zhou H, Sun K (2014) Multi-objective optimal planning of electric vehicle charging stations considering carbon emission. *Autom Electric Power Syst* 38(17):49–53
22. Xie ZC, Lin DH, Wan XS, Zhu WC, Lin TB, Chen LX (2016) A model of calculating the charging demands of electric vehicles in running based on large sample size of highway transportation data. *DESTech Trans Eng Technol Res (ICEEA)*
23. Zhong J, He L, Li C, Cao Y, Wang J, Fang B, Xiao G (2014) Coordinated control for large-scale EV charging facilities and energy storage devices participating in frequency regulation. *Appl Energy* 123:253–262
24. Xu H, Miao S, Zhang C, Shi D (2013) Optimal placement of charging infrastructures for large-scale integration of pure electric vehicles into grid. *Int J Electr Power Energy Syst* 53:159–165
25. Dong J, Liu C, Lin Z (2014) Charging infrastructure planning for promoting battery electric vehicles: an activity-based approach using multiday travel data. *Transp Res Part C Emerg Technol* 38:44–55
26. Liu P, Liu R, Bai X, Guo Z (2012) Charging load model based on diffusion theory for electric vehicles. *Electric Power Autom Equip* 32(9):30–34
27. Ip A, Fong S, Liu E (2010) Optimization for allocating BEV recharging stations in urban areas by using hierarchical clustering. In: 2010 6th international conference on advanced information management and service (IMS), pp 460–465. IEEE
28. Leung KK, Massey WA, Whitt W (1994) Traffic models for wireless communication networks. *IEEE J Sel Areas Commun* 12(8):1353–1364
29. Massey WA, Whitt W (1994) A stochastic model to capture space and time dynamics in wireless communication systems. *Probab Eng Inf Sci* 8(4):541–569
30. Bae S, Kwasinski A (2011) Spatial and temporal model of electric vehicle charging demand. *IEEE Trans Smart Grid* 3(1):394–403
31. Vibrations HRMMM (1977) Population dynamics, and traffic flow
32. Benes VE (2017) General stochastic processes in the theory of queues. Courier Dover Publications
33. Wu D, Aliprantis DC, Gkritza K (2010) Electric energy and power consumption by light-duty plug-in electric vehicles. *IEEE Trans Power Syst* 26(2):738–746

Research on the Time Interval of Discrete Traffic Flow Model



Liang Rui, Shan Zheng, and An Zeping

Abstract In order to verify the accuracy of the discrete traffic flow model and find out the most suitable statistical method of data for model validation, the discrete traffic flow model proposed by Robertson which was suitable for urban arterial between the intersections was studied. The traffic flow of five roads in Beijing city was investigated by multi-spot photography. The traffic flow data were collected at different time intervals according to 1–6 s and verify the discrete traffic flow model proposed by Robertson at different statistical intervals. Finally, the error of the model in different statistical interval of 1–36 s was obtained. The mean and variance of the error were analyzed and the optimal time interval of the survey data in this model was 1–5 s. This conclusion was verified by T test.

Keyword Traffic control · Platoon dispersion · Optimal time interval · Signal optimization

1 Introduction

With the accelerated process of urbanization and the rapid development of social economy, people's living standards were generally improved, the sharp increase in motor vehicles and traffic demand increases, "traffic jam" has become a common problem in modern cities. Serious traffic congestion will lead to additional travel time, increased vehicle energy consumption, air pollution and the increasing degree of noise pollution and many other issues. Urban traffic congestion problem was increasing year by year and the urgent need to be treated. The traffic control which is an effective means has the advantages of improving traffic order and safety. By means of traffic control, the traffic flow on the road can be guided and controlled reasonably to reduce the traffic accidents, noise pollution, exhaust emissions and ease traffic congestion. The theory of signal control in the intersection of the modern city shows that when the signal was matched according to the queuing vehicle, the

L. Rui (✉) · S. Zheng · A. Zeping
China Highway Engineering Consultants Corporation, Beijing 10089, China

traffic flow can be met the green light signal at the intersection without stopping to wait. The two-way traffic flow “green wave” formed will greatly improve the traffic congestion.

Therefore, this paper takes the traffic flow in the urban arterial as the research object and carries on the research to the acquisition and processing of the data on the basis of the traffic characteristics of the vehicle between the intersections. The movement of the vehicles in the queue at the intersections would be obtained. This research has certain practical significance for the control and optimization of urban arterial traffic signal.

2 Literature Review

Vehicles departing from the stop-line of an upstream signal intersection generally form a vehicle fleet. As vehicle speeds fluctuate during the motion, the fleet will disperse before it arrives at the downstream intersection, which was called “platoon dispersion”. Platoon dispersion is an essential phenomenon to be studied for building traffic models in traffic control system. Discrete traffic flow models directly influence the optimization of signal timing schemes in arterial signal coordination because of their impact on traffic delay models. Over the years, many scholars have studied the Discrete traffic flow model.

Pacey [1] proposed the traffic diffusion model assuming normal distribution of speed. Grace and Potts (1964) further investigated Pacey’s model from the density aspect. Later, Robertson [2] developed a recurrent dispersion model with field data gathered by Hillier and Rothery (1967), which was widely implemented in various signal plan design software and signal control systems. Seddon [3] reported that Robertson’s model was equivalently based on shifted geometric distribution of travel time. In 1990, Gantz and Mekemson [4] calibrated the parameters of the discrete model of the fleet and used the simulation program to compare the model before and after calibration. Cantarella [5] compared the discrete traffic flow model and the cell transmission model by studying the effects of the variation of the distances between pairs of adjacent junctions.

In China, the application and evaluation of the discrete traffic flow model of foreign countries was the focus of domestic research. Wei et al. [6] proposed a discrete traffic flow model for cars from the aspect of density with truncated normal distribution of speed assumption. Weiwei et al. [7] uses Vissim simulation software to simulate the line control system. By comparing the fleet discrete data before and after the change of road length, the relationship between the vehicle flow distribution and the fleet discrete characteristics is obtained. Yizhou et al. [8] use micro discrete model to study the mutual interference mechanism between people and vehicles at signalized intersections. Zhihong and Yangsheng [9], Zhihong et al. [10] proposed the discrete model of heterogeneous traffic flow fleet and the dynamic heterogeneous traffic flow fleet discrete model. Both models reduce the mean square root error of the prediction of the distribution of the arriving vehicle flow at the downstream intersection.

It can be seen that many scholars have mainly focused on the establishment of a new discrete traffic flow model or the analysis and improvement of existing discrete traffic flow model. With the progress of modern science and technology and data acquisition device, traffic data could be collected even by seconds. The requirements of data acquisition and processing were also getting higher and higher in the analysis and research of the discrete traffic flow model. However, literatures were lack of the research on the precision of data acquisition and processing. In view of this background, this paper takes the relevant data of the discrete traffic flow model as the starting point. What kind of time interval is suitable to study the statistics of the survey data? Because the geometric distribution model of Robertson was suitable for short distance travel time distribution and for the situation of the city between the two intersections, this paper chose the Robertson model for research.

3 Introduction to the Discrete Traffic Flow Model

In 1969, Robertson proposed a geometric distribution model. He believes that the actual travel time of the vehicle which between a fixed start and stop was distributed according to the geometric function. The dispersion coefficient of the fleet reflects the difference of travel speed. According to this analysis, the vehicle arrival rate of a certain section of the downstream and the vehicle pass rate upstream stop line section have the following relationship:

$$q_d(i + t) = Fq_0(i) + (1 - F)q_d(i + t - 1)$$

where

$q_d(i + t)$ —the expected rate of arrival of vehicles on a certain section of the downstream in the $(i + t)$ period.

$q_0(i)$ —the vehicle passing rate of upstream stop line section in the (i) period.

t —The average running time of the vehicle between the above two sections was 0.8 times (take the period number as a unit);

F —A coefficient which represents the discrete degree of traffic flow in the course of motion was called the discrete coefficient of traffic flow;

Robertson recommends that the formula for calculating F was:

$$F = \frac{1}{1 + A * t}$$

A : according to the observed values of the correction, A is usually taken 0.35; the other coefficients were the same as in the front.

Time interval T was a concept which was proposed to verify the accuracy of the Robertson geometric distribution model. The passing rate and the arrival rate of the geometric distribution model need to be obtained in the field investigation. Time

interval T was refers to use different time intervals to collect the data of the passing rate and the arrival rate. Since the rapid development of intelligent traffic control system in recent years, the accuracy of data acquisition device in traffic control has been calculated according to the second. Due to the time interval should not be more than the green light, the time interval range was set to 1–36 s. Then the model was verified separately.

4 Data Collection

The study of traffic characteristics can not be separated from the field survey of Road intersection data. In the study of the dispersion characteristics of the fleet, the field observation data is the key data of the later analysis. Through the field observation and analysis to master the basic change of traffic characteristics, this survey selected five sections of Beijing city. The specific circumstances and video conditions of the five sections is shown in Table 1.

Vehicles through the upstream intersection and vehicles arriving at the intersection of the downstream on the same road were recorded by a video camera. Since the camera has the function of recording time, the flow data of the four sections from the upstream and downstream was obtained from the video. Take the Sect. 1 as an example to introduce the collection and collation of survey data. In the case of the time interval was 1–6 s, the geometric distribution model of Robertson was used to calculate the field survey data. Then the error is calculated. Part of the Survey data and calculated data of Sect. 1 were shown in Table 2.

5 Study of the Time Interval

In order to verify the accuracy of the geometric distribution model of Robertson and get the optimal time interval, the error of the calculated data and the actual data was analyzed mathematically in the case of time interval T was 1–36 s. Finally, the mean value of each set of error data of the four sections was shown in Fig. 1.

By the analysis of the average value of the error, we can see that the average value of the error of each section was gradually increased with the increase of the time interval T . When the time interval $T = 1$ s, the average value of the error of the Section 1 was 0.55, the average value of the error of the Section 2 was 0.51, the average value of the error of the Section 3 was 0.58, the average value of the error of Section 4 was 0.66. When the time interval $T = 5$ s, the average value of the error of the Section 1 was 1.16, the average value of the error of the Section 2 was 1.2, the average value of the error of the Section 3 was 1.17, the average value of the error of Section 4 was 1.45. When the time interval $T = 36$ s, the average value of the error of the Section 1 was 13.4, the average value of the error of the Section 2 was 9.71, the average value of the error of the Section 3 was 10.01, the average value of the error

Table 1 The situation of survey

Section		Name	Phase of a signal light		
Section 1	Upstream intersection	The intersection of west road and north industrial road	Two phase	One-way lane number: Four lane	Distance of road (m): 553
	Downstream intersection	The intersection of west road and Ping Le Yuan district	Two phase	One-way vehicle flow (pcu/h): 1645	Video time: 9:30–10:30 am
Section 2	Upstream intersection	The intersection of Song Yu south road and west road	Four phase	One-way lane number: Four lane	Distance of road (m): 822
	Downstream intersection	The intersection of Song Yu south road and Wu Sheng road	Two phase	One-way vehicle flow (pcu/h): 1435	Video time: 10:30–11:30 am
Section 3	Upstream intersection	The intersection of south mill road and west road	Four phase	One-way lane number: Three lane	Distance of road (m): 803
	Downstream intersection	The intersection of south mill road and Wu Sheng road	Two phase	One-way vehicle flow (pcu/h): 1551	Video time: 2:30–3:30 pm
Section 4	Upstream intersection	The intersection of Song Yu north road and Wu Sheng road	Two phase	One-way lane number: Three lane	Distance of road (m): 794
	Downstream intersection	The intersection of Song Yu north road and west road	Two phase	One-way vehicle flow (pcu/h): 943	Video time: 3:30–4:30 pm

of Section 4 was 13.97. The average error values of all 4 arterials are calculated by 5 s interval and plotted with black bold lines in Fig. 1. The following data are shown in Table 3.

As can be seen from Table 6, the average value of the error of the four sections were less than 1 vehicles when the time interval $T = 1-5$ s. With the increase of the time interval T , the mean value of the error gradually increased to 3 vehicles, 7 vehicles. When the time interval $T = 30-36$ s, the mean value of the error has been increased to 10 vehicles.

This shows that the smaller the time interval T , the smaller the error between the actual data and the data calculated by the Discrete traffic flow model of Robertson, the more favorable to the accuracy of validation of the Discrete traffic flow model.

Table 2 Survey data and calculated data of Sect. 1

Time interval T (s)	1	0	1	0	1	0	1	2	3	0	2	0	2	1
1	Vehicle pass rate ($q_0(t)$)/pcu	1	0	0	0	0	0	0	0	0	0	0	0	0
	Vehicle arrival rate ($q_d(t + t - 1)$)/pcu	0	0	0	0	0	0	0	0	0	0	0	0	0
	Expected vehicle arrival rate ($q_d(t + t)$)/pcu	0.51	0.51	0	0	0.51	1.53	1.02	0	1.02	0	1.02	0	1.02
	Pass rate error /pcu	0.51	0.51	0	0	0.51	1.53	1.02	0	1.02	0	1.02	0	1.02
2	Vehicle pass rate ($q_0(t)$)/pcu	2	1	5	2	0	0	2	0	0	2	0	5	0
	Vehicle arrival rate ($q_d(t + t - 1)$)/pcu	0	0	0	0	0	2	2	0	1	1	1	1	0
	Expected vehicle arrival rate ($q_d(t + t)$)/pcu	1.02	0.51	2.55	1.02	0.98	2	0.49	0.98	2	0.49	3.04	0.49	3.04
	Pass rate error /pcu	1.02	0.51	2.55	1.02	0.98	2	0.49	0.98	2	0.49	3.04	0.49	3.04
3	Vehicle pass rate ($q_0(t)$)/pcu	2	6	2	2	2	6	3	2	3	6	3	1	1
	Vehicle arrival rate ($q_d(t + t - 1)$)/pcu	0	0	1	3	2	2	2	2	3.5	2	2.5	2.5	2.5
	Expected vehicle arrival rate ($q_d(t + t)$)/pcu	1.02	3.06	1.51	2.49	2	3.53	3.24	2	3.53	3.24	1.73	3.24	1.73
	Pass rate error /pcu	1.02	2.06	1.49	0.49	0	0.03	0.74	0	0.03	0.74	1.77	0.74	1.77

(continued)

Table 2 (continued)

Time interval T (s)																						
4	Vehicle pass rate $(q_0(t))/pcu$	3	7	2	5	5	1	3	2												
	Vehicle arrival rate $(q_d(i + t - 1))/pcu$	0	0	4	2	4.5	3.5	5.5	4												
	Expected vehicle arrival rate $(q_d(i + t))/pcu$	1.53	3.57	2.98	3.53	4.75	2.22	4.22	2.98												
	Pass rate error /pcu	1.53	0.43	0.98	0.97	1.26	3.28	0.22	0.02												
5	Vehicle pass rate $(q_0(t))/pcu$	6	4	4	8	2	2	2	3												
	Vehicle arrival rate $(q_d(i + t - 1))/pcu$	0	2	4	4.5	4.5	7.5	4	4												
	Expected vehicle arrival rate $(q_d(i + t))/pcu$	0	4.04	4	4.24	6.29	4.69	2.98	2.98												
	Pass rate error /pcu	2	0.04	0.5	0.26	1.21	0.69	1.02	1.98												
.....
36	Vehicle pass rate $(q_0(t))/pcu$	28	9.5	3	25	21	7	18	40												
	Vehicle arrival rate $(q_d(i + t - 1))/pcu$	26.5	14.5	3	13	24	10.5	11	32												
	Expected vehicle arrival rate $(q_d(i + t))/pcu$	27.27	11.95	3	19.12	22.47	8.71	14.57	36.08												
	Pass rate error /pcu	12.80	8.95	10	4.88	11.98	2.29	17.43	18.08												

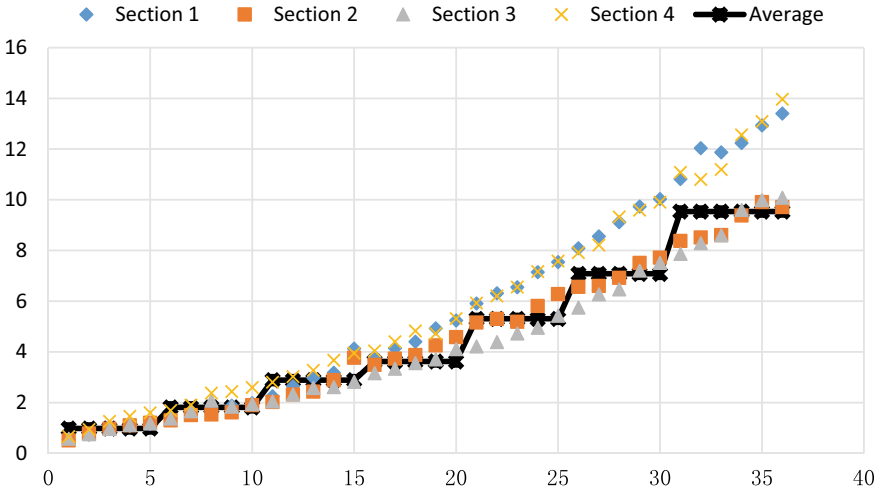


Fig. 1 Scatter plots of the mean value of the error

Table 3 Analysis of the average of the error

Time interval T (s)	1–5	5–10	10–15	15–20	20–25	25–30	30–36
Average of error	0.87	1.81	2.89	3.62	5.3	7.08	9.53

In order to make a further study on the volatility of the error between the survey data and the calculated data, this paper makes a mathematical analysis of the variance of the errors in each group. The results of the analysis were shown in Fig. 2.

In the analysis of the variance of each group, the trend of the variance of each group was gradually increasing with the increase of time interval T. When the time interval $T = 1$ s, the variance of the error of Section 1 was 0.4, the variance of the error of Section 2 was 0.38, the variance of the error of Section 3 was 0.5 the variance of the error of Section 4 was 0.01. When the time interval $T = 5$ s, the variance of the error of Section 1 was 1.14, the variance of the error of Section 2 was 1.31, the variance of the error of Section 3 was 1.46, the variance of the error of Section 4 was 2.3. When the time interval $T = 36$ s, the variance of the error of Section 1 was 59.52, the variance of the error of Section 2 was 50.24, the variance of the error of Section 3 was 42.08, the variance of the error of Section 4 was 51.23. In order to reflect the change of variance more clearly, the variance of the four sections will be calculated again according to the time interval interval of each 5 s. The average value of the graphic representation with black bold lines in Fig. 2. The data were shown in Table 4.

As can be seen from Table 4, the average value of the variance of the four sections was 0.32 when the time interval $T = 1-5$ s. This shows that the error of the calculated data and the actual data was very stable and each error value was close to the average value of the error when the time interval was 1–5 s. In the previous analysis of the

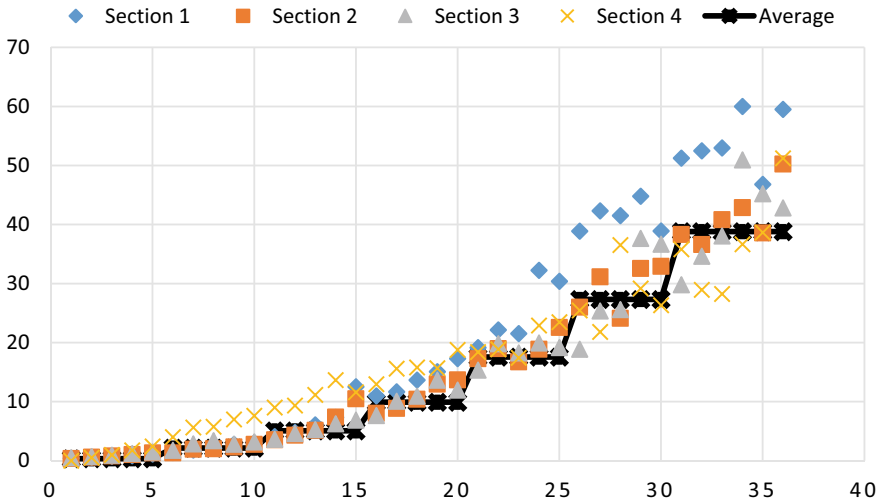


Fig. 2 Scatter plots of analysis of variance of error

Table 4 Analysis of the average of the variance

Time interval T (s)	1–5	5–10	10–15	15–20	20–25	25–30	30–36
Average of variance	0.32	2.13	5.06	9.89	17.55	27.3	38.8

average value, the average value of the error was less than 1 vehicles, that was the value of each error was less than 1 vehicles. The average value of the variance of the four sections increased to 38.8 when the time interval $T = 30-36$ s. This means that when the time interval was $30-36$ s, the fluctuation of the error of the calculated data and the actual data was very large. The degree of dispersion of each error value was very large, and the degree of the deviation from the mean of the error was different. Therefore, the data obtained in the case of time interval $T = 30-36$ s was not suitable for the Discrete traffic flow model to verify the accuracy.

Through the analysis of the error of the data of each group, it was not difficult to get the time interval T . With value smaller, the accuracy of the model validation was higher. Because the data of the four sections of the acquisition time were not peak hours, so the error is less than 1 vehicle to be considered a high precision verification. As a result, this paper obtains the time interval $T = 1-5$ s which was the best time to validate the team’s discrete model.

6 Verification of the Conclusion

For the conclusion of the best time interval T was 1–5 s, this paper uses the newly collected data to verify the conclusion of the Sect. 5. The information of Sect. 5 is shown in Table 5.

Processing method for data of Sect. 5 was the same as the processing method for the data of the previous four sections. The survey data of Sect. 5 were statistically analyzed according to the time interval of 1–5 s. Then the Statistical data would be calculated by the Discrete traffic flow model of Robertson. Finally, five sets of vehicle arrival rates at the downstream of different time interval were obtained. Part of the survey data and calculation data of Sect. 5 were shown in Table 6.

In this paper, we use the T test in mathematical statistics to verify the above conclusions. T test was a test method which was used to determine whether the difference between the average of the two groups was significant. T test was suitable for the general distribution of normal distribution which was small samples and the standard deviation of the unknown. Therefore, we can use T distribution theory to infer the probability of the occurrence of the difference, and then determine whether the difference between the two average was significant.

The calculation data and the actual data of time interval $T = 1-5$ s were imported into SPSS software. The output results obtained by using the T test were shown in Table 7.

The Levene test of the variance equation was also called the homogeneity test of variance. From the above table shows that when the time interval $T = 1-5$ s, the sig value of the levene test of the variance equation is greater than 0.01. This shows that there was no significant difference between the calculated data and the actual data of the five groups of data. The sig value (P value) of the T test of mean values were greater than 5%, that is, there was no significant difference between the calculated data and the actual data in the five groups. Therefore, the data can be used to verify the accuracy of the Discrete traffic flow model accurately when the time interval $T = 1-5$ s, which verifies the conclusion that the $T = 1-5$ s was the best time interval.

Table 5 The situation of Sect. 5

	Name	Phase of a signal light		
Upstream intersection	The intersection of west road and south mill road	Four phase	One-way lane number: Four lane	Distance of road (m): 1000
Downstream intersection	The intersection of west road and Guang Qu road	Four phase	One-way vehicle flow (pcu/h): 1226	Video time: 10:30–11:30 am

Table 6 Survey data and calculation data of Sect. 5

Time interval T (s)																					
1	Vehicle pass rate $(q_0(t))/pcu$	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Vehicle arrival rate $(q_d(i + t - 1))/pcu$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Expected vehicle arrival rate $(q_d(i + t))/pcu$	0.46	0.23	0.2	0.2	0	0	0.23	0	0	0.23	0	0	0	0	0	0	0	0	0
2	Vehicle pass rate $(q_0(t))/pcu$	1	0	1	2	2	1	1	1	1	1	0	0	0	0	0	0	0	0	0
	Vehicle arrival rate $(q_d(i + t - 1))/pcu$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Expected vehicle arrival rate $(q_d(i + t))/pcu$	0.69	0.23	0.2	0.2	0	0	0.23	0	0.2	1.23	0.5	1.77
3	Vehicle pass rate $(q_0(t))/pcu$	2	3	1	0	2	2	2	2	2	3	1
	Vehicle arrival rate $(q_d(i + t - 1))/pcu$	0	1	2	1	5	2	2	2	3	1
	Expected vehicle arrival rate $(q_d(i + t))/pcu$	0.23	0.23	1.2	2.23	1	3.85	2	2.77
4	Vehicle pass rate $(q_0(t))/pcu$	3	1	2	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Vehicle arrival rate $(q_d(i + t - 1))/pcu$	1	2	6	2	3	4	1	3	4	1	3

(continued)

Table 6 (continued)

Time interval T (s)													
5	Expected vehicle arrival rate($q_d(i+t)$)/pcu	0.23	1.45	2.2	4.85	2	3.23	3.5	0.77			
	Vehicle pass rate ($q_0(i)$)/pcu	5	1	2	5	1	7.5	4.5	2			
	Vehicle arrival rate ($q_d(i+t-1)$)/pcu	1	2	7	4	4	1	3	4			
	Expected vehicle arrival rate($q_d(i+t)$)/pcu	0.92	1.23	2.7	5.62	3.5	4.23	1	4.03			

Table 7 Independent sample test

Time interval T (s)	Levene test of variance equation	T test for mean value equation						
		F	Sig	T	df	Sig(bilateral)	Mean difference	Standard error value
1	Assume that the variance was equal	5.663	0.018	-0.13	198	0.895	-0.01	0.087
	Assume that the variance was not equal			-0.13	189.8	0.895	-0.01	0.087
2	Assume that the variance was equal	2.999	0.085	- 0.08	198	0.939	-0.01	0.136
	Assume that the variance was not equal			-0.08	192.8	0.939	-0.01	0.136
3	Assume that the variance was equal	1.854	0.175	0.066	198	0.948	0.011	0.17
	Assume that the variance was not equal			0.066	192.2	0.948	0.011	0.17
4	Assume that the variance was equal	2.107	0.148	-0.01	198	0.991	-0	0.196
	Assume that the variance was not equal			-0.01	194.8	0.991	-0	0.196

(continued)

Table 7 (continued)

Time interval T (s)	Levene test of variance equation	T test for mean value equation						
		F	Sig	T	df	Sig(bilateral)	Mean difference	Standard error value
5	Assume that the variance was equal	6.14	0.013	-0.21	1178	0.834	-0.02	0.098
	Assume that the variance was not equal			-0.21	1162	0.834	-0.02	0.098

7 Concluding Remarks

In this paper, the traffic flow of five roads in Beijing city was investigated by multi-spot photography. The traffic flow data were collected at different time intervals according to 1–36 s and verify the discrete traffic flow model proposed by Robertson at different statistical intervals. Finally, the error of the model in different statistical interval of 1–36 s was obtained. Through the mathematical analysis of the mean and variance of the error, the conclusion was drawn that the arrival rate and pass rate of the model was 1–5 s, then the T test method was used to verify the conclusion. This conclusion provides the most direct method for the statistics and processing of traffic survey data in the future. To a certain extent, this conclusion has a positive effect on the research of the discrete phenomena of traffic flow and the optimization of signal timing at the intersection.

References

1. Pacey GM (1956) The process of a bunch of vehicles released from a traffic signal, RRL report LR 253. Road Research Laboratory Research Note No. RN/2665/GMP
2. Robertson DI (1969) TRANSYT—a traffic network study tool, RRL report LR 253. Road Research Laboratory, Crowthorne, Berkshire, UK
3. Seddon PA (1972) Another look at platoon dispersion: 3. The recurrence relationship. *Traffic Eng Control* 13(10):442–444
4. Gantz DT, Mekemson JR (1990) Flow profile comparison of a microscopic car-following model and a macroscopic platoon dispersion model for traffic simulation. *Simul Conf ISPN: 770–774*
5. Cantarella GE (2013) The network signal setting problem: the coordination approach vs. the synchronisation approach. In: *Computer modelling and simulation (UKSim)*, ISBN: 575–579

6. Wei M, Jin W, Shen L (2012) A platoon dispersion model based on a truncated normal distribution of speed. *J Appl Math* 727839:13–13
7. Weiwei S, Kang L, Binbin H (2015) Research on fleet discrete characteristics based on VISSIM. *Transp Technol Economy* 3:43–48
8. Yizhou T, Yanfang W, Qingfei G, Liyun D (2019) Pedestrian-vehicle interference at a signalized crossing based on detailed microscopic traffic flow models. *J Phys* 68(24):102–114
9. Zhihong Y, Yangsheng J (2018) Dynamic Robertson fleet discrete model in the environment of Internet of vehicles. *J Southwest Jiaotong Univ* 53(02):385–391
10. Zhihong Y, Yangsheng J, Yi W, Yanru C (2019) Discrete model of dynamic heterogeneous traffic flow fleet under the environment of Internet of vehicles. *J Beijing Jiaotong Univ* 43(02):107–116

An Approach to Analyze Commuters' Transfer Characteristics on the Public Transportation Network in an Incomplete Data Environment



Shichao Sun

Abstract This paper analyzed commuters' transfer characteristics on public transportation networks in an incomplete data environment where only smartcard data could be accessed. A frequent-patterns-mining method was proposed, and a case study was given in Shenzhen, China. Specifically, commuters were firstly identified by using smartcard data based on their bus-riding regularity on workdays. Then, frequent correlations between bus routes/metro lines used in long-term rush hours were measured for each commuter. A two-level minimum threshold framework was proposed to evaluate the correlations, and then determine whether commuting trips containing transfers. The results indicated that non-transfer commuting trips accounted for nearly 72% of the total, among which about 87% were taking buses. Moreover, 14% of identified commuters were having commuting trips through one transfer, where interchanges within the same transportation mode were the majority, such as transfers from bus to bus, and transfers between metro lines. It reflected that the efficiency and convenience of transfers from bus to metro needed to be improved.

Keywords Transfer characteristics · Smartcard data · Frequent patterns mining

1 Introduction

Public transportation (PT) has been widely regarded as an effective way to alleviate serious traffic congestion in metropolises. Its effect is very obvious especially in China, where the private car ownership has gone through a rapid increase in the last decade [1–3]. Hence, the local government has been working for years to promote the development of bus priority. Nevertheless, in order to achieve a better sustainable public transportation system, it not only relies on the new construction of infrastructures on the supply-side, but also requires an improvement of high-quality service to meet with passengers' increasing travel demands. In this context, a better understanding of the efficiency and service quality related to transfers could promote the

S. Sun (✉)
Dalian Maritime University, Dalian, China
e-mail: dlmu_sunshichao@163.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_3

37

competitiveness of the public transportation system in a big picture. Therefore, a high efficiency of interchanges within public transportation systems were regarded as a key part to the success, and the transfer pattern analysis has received much attention. In this regards, a number of studies have provided insights into commuters' transfer characteristics in rush hours [4–6]. The accurate knowledge of their transfer behaviors was of great significance, because these time-rigid users were more demanding on the efficiency and convenience of transfers. However, it is challenging to collect commuters' detailed travel behaviors in both temporal and spatial dimensions on the public transportation network. Survey data are commonly used to achieve depicting their travel behaviors, but it is often plagued by its high cost, low data accuracy and low response rate. Conversely, Automatic Fare Collection (AFC) systems could provide more abundant, less-cost and high-quality smartcard data to take place of survey data in some extent [7]. The data generated from individuals' smartcard usage could be adopted to reflect their dynamic travel behaviors in a more fine granularity.

Therefore, this paper aimed at mining smartcard data to analyze public transportation commuters' transfer characteristics. In this respect, two main problems were involved:

- (a) **How to identify commuters in smartcard data?**
- (b) **How to distinguish “true” transfers (users taking the next route immediately after finishing the previous one) from “false” transfers according to their travel behaviors derived from smartcard data?**

As for the first problem, it is not a straightforward task to identify commuters from cardholders, because smartcard data were usually anonymous [8]. A number of approaches related to clustering methods were commonly proposed to separate commuters and non-commuters in cardholders, and spatio-temporal regularities of cardholder's travel behaviors were usually employed as primary indicators [4, 5, 9–12]. Nevertheless, it still remains unclear that what kinds of characterized travel patterns tied to public transportation commuters [4]. On the other hand, it was also necessary to identify transfer behaviors in the entire trips, so that they can be incorporated into future analyzes. However, individuals' transactions generated from each trip were not automatically attached with a transfer mark in smartcard data. Thus, it is still hard to directly reveal “true” transfers according to raw smartcard data. To this end, a number of studies enrich the dataset by integrating smartcard data with bus location information harvested from on-board AVL systems [11, 13, 14]. As a consequence, bus boarding stations as well as alighting stations corresponding to each smartcard record could be estimated. As for transactions in metro systems, information of origin and destination of each trip were directly attached to each smartcard record. Therefore, geometry relationships (distances, etc.) between previous alighting points and the next boarding location in two consecutive trips could be inferred, and corresponding potential time–cost including walking time and waiting time could be estimated. Then, a time threshold was usually set up to determine transfer behaviors by examining whether it matched with the potential time–cost [4–6, 9, 11, 15–20].

However, the fusion of smartcard data and AVL data was the main data source that supported the implementation of above approaches to identify commuters and “true” transfers. Unfortunately, in many developing countries such as China, not all buses have been equipped with both AFC facilities and on-board AVL equipment. Specifically, AFC facilities have been widely equipped in decades to take place of manual fare collections. On the contrary, the equipment ratio of on-board AVL devices was relatively low. Taking Shenzhen as an example, it was one of the most popular mega cities in China. Up to 2017, only 230 out of total 977 bus lines (nearly 6000 buses out of total 20,000 buses), have been equipped with both AFC facilities and independent AVL systems. The situation in other cities is even worse. Thus, other than above conventional approaches, how to analyze commuters' transfer characteristics in such an incomplete data environment was the key in the current study. The main contributions of this paper were:

- Public transportation commuters were identified from cardholders based on their behavioral regularities.
- Since AVL data were absent, we addressed transfer identification by mining frequent patterns of commuters' recurring travel behaviors in rush hours. Measures of travel behaviors were derived from sole smartcard data.

The rest of paper was organized as follows: In Sect. 2, smartcard data harvested from Shenzhen, China were described. Subsequently, public transportation commuters were identified based on the analysis of their behavioral regularities derived from smartcard data. Then, data mining tools were introduced to capture “true” transfers in the incomplete data environment, and transfer characteristics of the identified commuters were analyzed and discussed in Sect. 4. In Sect. 5, conclusions were drawn and future directions were given.

2 Data Preparation

Up to 2017, there are 8 metro lines and over 950 bus routes served in Shenzhen's urban public transportation system. According to Shenzhen transport annual report in 2017, approximately 6 million residents used public transportation daily. “Shenzhen Tong” was the only contactless smartcard used for electronic payments in the public transportation system. Passengers can access to all bus routes and metro lines through “Shenzhen Tong”, and over 95% of travel transactions were completed through AFC systems in Shenzhen.

Smartcard data analyzed in this paper were harvested from “Shenzhen Tong”. The original dataset contained entire transaction records on the public transportation network during from 3rd to 30th, July, 2017. Thus, 20 workdays and 3 weekends were covered in this time period. Each record contains several fields such as card id (unique primary key), transaction date, timestamp, event type, and route information (Table 1). Specifically, card-ids initially with “05”, “15” or “25” are registered by students, while rest cards are regular. In addition, event type marked with “21” and

Table 1 Examples of smartcard data

Card id	Date	Transaction time	Event	Bus route/metro line
575,483,827	2017/7/14	11:14:54	21	Metro line 1
253,958,383	2017/7/10	12:55:23	22	Metro line 4
593,832,848	2017/7/24	15:23:34	31	Bus route 301
153,948,574	2017/7/12	06:45:53	22	Metro line 3
948,372,344	2017/7/13	07:23:45	31	Bus route 512

“22” respectively represent inbound and outbound at metro stations, while event “31” represent swiping cards on buses. Bus fare is not distance-based pricing in Shenzhen, so passengers only need to swipe their card when boarding. Moreover, passengers also do not need to swipe their cards again for the interchanges between two metro lines. That is, smartcard data will only record the metro line information of passengers’ inbound and outbound stations.

3 Methods

3.1 Definition of Peak-Hours

Ridership of a public transportation system on workdays will be typically fluctuate according to the time of day. Generally, a typical ridership curve at the time of day will demonstrate two obvious ridership-peaks, which respectively corresponds to morning and evening rush hours. Thus, in order to measure relatively precise peak-hours in Shenzhen, collected smartcard data were used to count the ridership at time of day (half-hour interval). Both bus boarding records and metro inbound records in 20 workdays were considered to obtain several temporal distributions (Fig. 1). The results indicated that the morning peak in Shenzhen basically started from 6:30 am and continued to almost 9:30 am. In addition, the second peak in the evening seemed to last longer than the previous one, and it ranged from 4:30–8:30 pm.

3.2 Identification of Commuters

It is assumed that commuters’ travel demands accounted for most of the ridership in peak hours, and their travel behaviors in commuting trips were generally recurring. Specifically, typical travel regularity of commuters on workdays could be denoted as taking buses/metro to the workplace from home in the morning and then returning home in the evening [6]. Therefore, passengers’ behavioural regularities within rush hours could be analysed, and characterized travel patterns could be adopted as rules

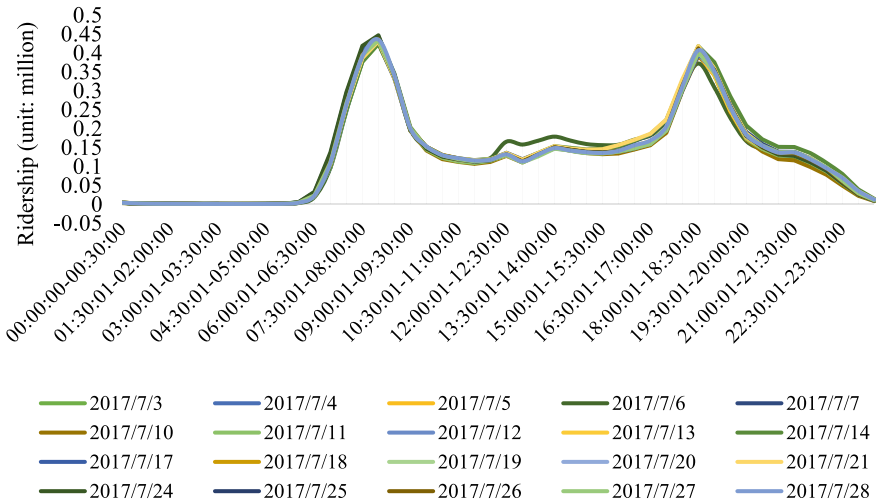


Fig.1 An illustration of the fluctuation of ridership at the time of day (20 workdays)

to separate commuters and non-commuters. The peak-hours were defined as above, which could be employed as a given time period for analysis. Passengers' travel behaviors during this time period could be observed by smartcard data, and characteristics of smartcard use in both temporal and spatial dimensions could be adopted to reflect their behavioral regularity [4]. In this study, we defined D_B as the number of workdays where cardholders swipe their cards in both morning peak and evening peak of one day (on a maximum of 20 days). Thus, D_B could be regarded as an indicator reflecting the degree of repetition of potential commuting trips, which could also be employed as the threshold to distinguish commuters [10]. However, it was still hard to determine the value of threshold, and also couldn't confirm the accuracy of the identification results.

In order to cope with this problem, the cardholder distribution was calculated based on D_B , shown in Fig. 2a. Moreover, the distribution of transaction records was illustrated in Fig. 2b, where each pillar represented the number of records generated in peak-hours by the cardholders of corresponding D_B in Fig. 2a. As a consequence, it reported that approximately 3 million cardholders were involved, and each of them had at least one workday where he/she swiped the card in both the morning peak and evening peak of that day. Furthermore, nearly a quarter of involved cardholders only had one workday satisfied the above condition. The number of cardholders corresponding to D_B less than 8 accounted for almost 60% of the total. Nevertheless, surprisingly, according to the counterpart in Fig. 2b, these cardholders only produced about 20% of total transaction records. On the contrary, the proportion of cardholders corresponded to D_B more than 14 was nearly 25%, but transaction records generated by these passengers accounted for more than half of the total.

Generally, the larger D_B is, the more likely corresponding cardholders will be "true" commuters. For instance, cardholders with D_B more than 15 may have a great

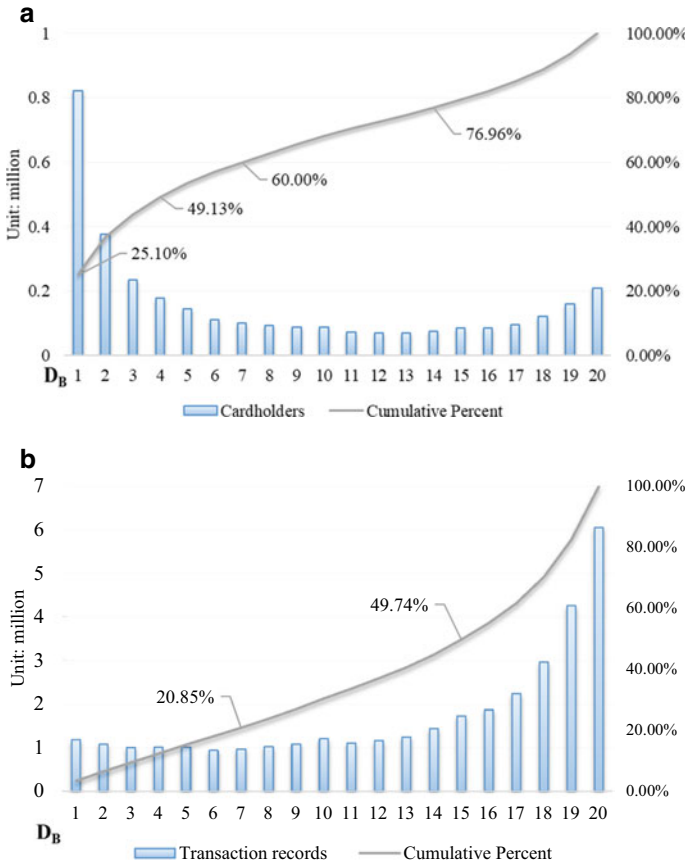


Fig. 2 **a** The distribution of cardholders based on D_B and, **b** the distribution of corresponding transactions based on D_B

probability to be commuters, because they actually satisfied such a characterized regularity condition for about three quarters of 20 workdays. Nevertheless, in order to better validate and improve the accuracy of identification, we firstly cross-checked the stability of bus route/metro lines which were used in rush hours by cardholders with D_B more than 11. Specifically, for workdays where the cardholder swiped his/her card in both peak-hours, a dataset was set up to contain all the routes used in all his/her first trips of these days. Then, stability of bus routes/metro lines corresponding to their first trips in PT systems was analyzed for each cardholder based on the dataset. It indicated that the bus routes/metro lines in the datasets were relatively stable (Table 2). When D_B increased to 14 or larger, more than 80% of the corresponding cardholders relied on the same bus route/metro line in all their first trips on workdays. On the contrary, obviously, stability declined rapidly when D_B decreased to 13 or lower. Additionally, we also checked whether the bus route/metro line corresponding to these cardholders'

Table 2 Statistical indicators of each group

Group	Stability of the 1st bus route/metro line	Route correlation between first and last trip
$D_B = 12$	65.6% of cardholders had no changes	87.7% of group members satisfied the condition
$D_B = 13$	72.9% of cardholders had no changes	89.5% of group members satisfied the condition
$D_B = 14$	83.6% of cardholders had no changes	97.4% of group members satisfied the condition
$D_B = 15$	87.1% of cardholders had no changes	95.6% of group members satisfied the condition
$D_B = 16$	89.3% of cardholders had no changes	96.4% of group members satisfied the condition
$D_B = 17$	90.8% of cardholders had no changes	97.9% of group members satisfied the condition
$D_B = 18$	89.7% of cardholders had no changes	96.2% of group members satisfied the condition
$D_B = 19$	88.1% of cardholders had no changes	98.7% of group members satisfied the condition
$D_B = 20$	87.9% of cardholders had no changes	98.6% of group members satisfied the condition

last trips per workday strongly correlated with the counterparts in their first trips. It is because a typical commuting chain reflected a home-to-workplace-to-home trip, the bus route/metro line used in commuter's last trip should have a strong correlation with the ones used in his/her first trips. Thus, it was assumed that if the bus route/metro line used in one cardholder's last trip was also contained in the route dataset of his/her first trips, then we could determine a strong correlation. The results indicated that more than 90% of involved cardholders satisfied the correlation condition when D_B was more than 13. Due to the overlap of bus routes, cardholders' selection of which bus routes to be used in their first and last trips may not be absolutely fixed from day to day. As a consequence, taking account of the behavioral regularity of commuters and the statistics calculated above, the threshold for the identification was finally set as 13. It implied that cardholders with D_B more than the threshold could be regarded as commuters, and approximately 690,000 commuters were identified. Taking 3rd, July 2017 as an example, identified commuters contributed more than 80% of transactions during morning peak-hours as well as evening peaks. Thus, analysis of transfer characteristics analysis on this group of passengers was sufficiently representative.

3.3 Identification of Transfers in Commuting Trips

How to identify interchanges in commuters' trips during rush hours was the key to better analyze their transfer patterns. The conventional way to achieve the identification in smartcard data environments is commonly setting up an appropriate time threshold to distinguish "true" transfers. Generally, the time threshold is defined as a potential time cost for transferring from the previous drop-off station to the next boarding station. Thus, it shall consider the walking time from the previous alighting station to the next boarding station, and also the waiting time at the next boarding station [10, 11]. Then, the actual time difference between two trips will be estimated by using smartcard data, and it will compare with the time threshold to determine if it satisfied transfer condition.

However, the time threshold should be setup based on different transfer scenarios. For instance, passengers' inbound and outbound time at metro stations could be precisely obtained by using smartcard data. On the contrary, as for bus systems, it is not yet a straightforward task to obtain passengers' alighting time, but it could be inferred by using the fusion of smartcard data and AVL data. Therefore, it was noted that the time threshold set up for the identification of interchanges from metro to bus (M2B) only need to consider the walking time between previous outbound metro station and the next bus boarding station, and also the waiting time at the next boarding station. However, with regard to the identification of interchanges from bus to other modes, such as bus to bus (B2B) and bus to metro (B2M), travel time of the previous bus trip should be additionally considered other than walking time and waiting time. It is because the alighting station in the previous bus trip was inferred, and the drop-off time needed to be estimated based on the travel time of that trip, which would influence the setup of corresponding time threshold. Nevertheless, it was still hard to determine a standardized time threshold for all candidate B2B or M2B transfers. The situation would become even worse especially in an incomplete data environment where AVL data were absent. In this context, both bus boarding stations and drop-off stations could not be inferred, which implied travel time of bus trips could not be estimated to achieve determining the time threshold. As for transfers within metro systems (M2M), it is not necessary to determine a time threshold, since passengers do not need to swipe their cards again for the interchanges between two metro lines. Thus, based on the information of passengers' inbound and outbound stations derived from smartcard data, we could determine if having interchanges or not by using the shortest path matrix between any two stations in the subway network.

Therefore, this paper intended to shed light on commuters' transfer characteristics on the public transportation system in such an incomplete data environment. Due to the lack of AVL data, we tried to address M2B, B2B and B2M transfer-identification problem through better understanding the characterized regularity of behaviors in recurring commuting trips. Thus, one main hypothesis of this study was that commuter's transfer behaviors would be recurring in their commuting trips during rush hours. Specifically, it implied that certain bus routes/metro lines would be frequently used in one's long-term commuting trips during peak hours. On this

basis, furthermore, we could also infer that if a commuting trip contained a transfer, it would demonstrate a strong correlation between the bus routes/metro lines used in the previous trip as well as the next trip in a long-term view. For example, one commuter may frequently “take Bus Route A firstly, and then transferred to Metro Line B” in his/her daily commuting trip. In other words, Bus Route A and Metro Line B are usually used together in a certain order during rush hours, and Metro Line B may have a great probability to be used following Bus Route A was taken. In this context, a Frequent Patterns Mining (FPM) method was employed here to identify transfers in commuting trips. It was a data mining technique widely applied in commercial areas. For instance in Wal-Mart, male customer’s frequently purchasing beer following their purchases of diapers would be of great help to shelf layout. With regards to the application of FPM in this study, bus routes/metro lines used orderly in one’s long-term commuting trips would be paid close attention to examine the frequent correlation between them.

In this paper, a case study in Shenzhen, China was given, and smartcard data produced by above identified commuters during morning rush hours (6:30–9:30 am) were incorporated into further analyzes. Firstly, commuting trips were extracted from the smartcard data for each commuter, which contained one’s orderly routes information used during morning peaks on each workday, as shown in Table 3. However, it should be noted that M2M transfers could be identified directly by using the shortest path matrix as well as inbound and outbound information derived from smartcard data. Therefore, passengers’ actual paths within the metro system could be revealed previously, and then incorporated into corresponding commuting trips. Then, the FPM approach could be applied to identify transfers in commuting trips based on the frequent correlation of one’s orderly bus routes/metro lines.

Table 3 Examples of one’s commuting trips derived from smartcard data

No.	Card ID	Date	1st boarding route	2nd boarding route	Route sequence
399	10183	2017/07/03	Bus route A	Metro line B	A → B
400	10183	2017/07/04	Bus route A	Metro line B	A → B
401	10183	2017/07/05	Bus route A	Metro line C	A → C
402	10183	2017/07/06	Bus route D	Metro line B	D → B
403	10183	2017/07/07	Bus route A	Metro line B	A → B
404	10183	2017/07/10	Bus route A	Metro line C	A → C
405	10183	2017/07/11	Bus route A	Metro line B	A → B
406	10183	2017/07/12	Bus route A	Metro line C	A → C
407	10183	2017/07/13	Bus route A	Metro line B	A → B
408	10183	2017/07/14	Bus route E	Metro line B	E → B
409	10184	2017/07/04	Metro line C	Metro line B	C → B
410	10184	2017/07/05	Metro line B	Metro line F	B → F
411	10184	2017/07/08	Metro line C	Metro line B	C → B
412	10184	2017/07/11	Metro line C	Metro line B	C → B

Generally, FPM could be represented in the form of association rules. Taking the cardholder “10183” in Table 3 as an example, he/she used *Bus Route A* \rightarrow *Metro Line B* in order frequently during morning peak-hours on workdays, so that the corresponding association rules could be denoted as:

$$\text{Bus Route A} \Rightarrow \text{Metro Line B} [\text{support} = 50\%, \text{confidence} = 62.5\%] \quad (1)$$

$$\text{Support} (\text{Bus Route A} \Rightarrow \text{Metro Line B}) = P(A \cup B) \quad (2)$$

$$\text{Confidence} (\text{Bus Route A} \Rightarrow \text{Metro Line B}) = P(B|A) \quad (3)$$

Specifically, the support degree and confidence degree were the two measures in the association rules, which respectively reflected the validity and certainty of the discovered frequent correlation between routes/lines. Support degree denoted in (2) was equal to 50% in the above case, because there are 5 out of total 10 commuting trips showing that the commuter took Bus Route A firstly, and then transferred to Metro Line B. On the other hand, the confidence degree was equal to 62.5%, which was calculated by using Bayesian probability in (3). It implied the probability of using Metro Line B following Bus Route A was taken. Moreover, the association rules in (1) were an example for the dataset including 2 boarding routes in order, namely frequent 2-routes-set. The same rules and measures are also applicable to frequent 3-routes-set, such as “*Bus Route A* \rightarrow *Metro Line B* \rightarrow *Metro Line A*”, namely two transfers in commuting trips. The standardized steps for frequent patterns mining through association rules were as follows:

Step 1: Find all frequent routes-sets, such as 2-routes-sets and 3-routes-sets.

In this context, a minimum support count should be pre-determined, which was used as a threshold to select candidate frequent routes-sets. That is, support degrees of each candidate routes-sets should be equal to or larger than the minimum support count. Then, the selected candidate routes-sets could be further analyzed in the next step to determine whether it contained transfers.

Step 2: Generate association rules of the frequent datasets for measuring.

Frequent patterns could be generated by mining the association rules of eligible routes-sets. Then, potential transfers in each routes-set could be confirmed, if its association rules satisfied both the minimum thresholds of support degree and confidence degree. These thresholds could be set by users or experts. For example, if the minimum threshold of measures were set as [support = 40%, confidence = 50%], then the case described above could be considered as a qualified frequent 2-routes-set. It was revealed that the commuter interchanged between the two routes in his/her commuting trips.

In addition, although association rules could be generated for the measure of correlations between bus routes/metro lines from the perspective of probability, it was also necessary to take account of the spatial geography relationship between them. That is, it should be determined whether there could be a potential transfer point in geospatial between the two routes. Therefore, in this study, we assumed that

if the walking distance between the nearest stations of the two routes was within an acceptable range (usually set to 500 m according to the service radius of a bus station in China), there could be a potential transfer point between them. Otherwise, the two bus routes/metro lines should be separated, and it seemed to be no potential transfers between them, regardless of how strong the frequent correlation between them was.

4 Findings and Discussions

Due to the increasing of routes overlaps in the public transportation system, both the support degree and confidence degree of potential transfers in the association rules will decrease significantly. It is because commuters may have more options of routes to fulfill their same travel demands. In this context, a two-level framework of association rules was employed as the measurement for transfer identification. Specifically, a low-level minimum threshold [support = 25%, confidence = 50%] was set to alleviate the influence of high overlap factor on the identification, and a high-level minimum threshold [support = 50%, confidence = 70%] was adopted as a tighter constraint to improve the reliability of results. However, since it was indeed hard to determine an exact threshold for the identification, the actual value was supposed to fall between the results of two levels. The qualified association rules may contain frequent 1-routes-set, frequent 2-route-set and frequent 3-route-set, which respectively represented non-transfer commuting trips, commuting trips through one transfer and commuting trips through two transfers. The identification results were shown in Fig. 3, where the distribution of commuters was illustrated according to their identified transfer type. With respect to results based on low-level thresholds, it indicated that commuters having non-transfer commuting trips



Fig. 3 Distribution of commuters according to identified transfer types

accounted for 75% of the total. At the same time, 17% of identified commuters were having one transfer in their commuting trips, and the proportion went down to 2% when considering commuting trips with two transfers. The rest (6%) of identified commuters had irregular travel patterns in their commuting trips, and none transfer type could be recognized. As for counterparts based on high-level thresholds, 72% of commuters were identified to have non-transfer commuting trips, 14% for one-transfer commuting trips, and 2% for two-transfer commuting trips. Thus, regardless of picking low-level thresholds and high-level threshold, there seemed to be only a slight difference between the results, which reflected a high reliability of the identification results.

Then, based on the identified transfers under the high-level measurement, bus routes/metro lines used in the commuting trips was analyzed. As for non-transfer trips, nearly 87% of the corresponding commuters relied on buses, and metro travels in commuting trips only accounted for 13%. With regards to commuting trips with one transfer, B2B transfers accounted for about 39% of the total, and the proportion of M2M transfers went up to 49%. Surprisingly, the proportion of B2M transfers was unexpectedly low (about 9%). As for commuting trips through two transfers, (B2M2M) transfers were dominated (67% of the corresponding commuting trips). The results indicated that the development of the common bus system in Shenzhen was relatively rapid and well, so commuters relied much on buses in their commuting trips. However, efficiency and convenience of transfers between bus stations and metro stations seemed to be insufficient, since B2M transfers unexpectedly only accounted for only 9% of the total commuting trips with one transfer.

5 Conclusions

This paper analyzed commuters' transfer characteristics on public transportation networks in an incomplete data environment where bus boarding and alighting locations were absent. A frequent-patterns-mining method was proposed, and a case study was given in Shenzhen, China. Specifically, commuters were identified based on their behavioral regularity of smartcard usage on workdays. Then, FPM method was employed for the identification of transfers in commuting trips by examining the frequent correlations between bus routes/metro lines. The results reflected a high reliability of identification results, based on which the transfer characteristics of commuters were analyzed. It indicated that:

- Non-transfer commuting trips accounted for nearly 72% of the total. At the same time, 14% of identified commuters were having commuting trips through one transfer, and commuting trips with two transfers were relatively rare (2%).
- Specifically, about 87% of the identified non-transfer commuters were relying on buses. As for commuting trips through one transfer, B2B transfers and M2M transfers were the majority. However, the proportion of B2M transfers was unexpectedly low. B2M2M transfers dominated in the identified corresponding commuting

trips. Nevertheless, more attempts should be also implemented to improve the accuracy of the identification of transfer journeys in further analyzes.

Acknowledgements This study is supported by “National Natural Science Foundation of China (71702019)”, and “the Fundamental Research Funds for the Central Universities (3132020163)”.

References

1. Sun S, Duan Z (2019) Modeling passengers' loyalty to public transit in a two-dimensional framework: a case study in Xiamen, China. *Transp Res Part A Policy Practice* 124:295–309
2. Sun SC, Xu LY, Yao YK, Duan ZY (2021) Investigating the determinants to retain spurious-loyalty passengers: A data-fusion based approach. *Transp Res Part A Policy and Practice* 152:70–83
3. Sun SC, Yang DY (2018) Identifying public transit commuters based on both the smartcard data and survey data: a case study in Xiamen, China. *J Adv Transp* 2018
4. Ma XL, Liu CC, Wen HM, Wang YP, Wu YJ (2017) Understanding commuting patterns using transit smart card data. *J Transp Geogr* 58:135–145
5. Zhou JP, Murphy E, Long Y (2014) Commuting efficiency in the Beijing metropolitan area: an exploration combining smartcard and travel survey data. *J Transp Geogr* 41:175–183
6. Ma XL, Wu YJ, Wang YH, Chen F, Liu JF (2013) Mining smart card data for transit riders' travel patterns. *Transp Res Part C Emerg Technol* 36:1–12
7. Pelletier MP, Trepanier M, Morency C (2011) Smart card data use in public transit: a literature review. *Transp Res Part C-Emerg Technol* 19(4):557–568
8. Kusakabe T, Asakura Y (2014) Behavioural data mining of transit smart card data: a data fusion approach. *Transp Res Part C Emerg Technol* 46:179–191
9. Chen J, Yang DY (2013) Identifying boarding stops of bus passengers with smart cards based on intelligent dispatching data. *J Transp Syst Eng Inf Technol* 13(1):76–80
10. Shi XY, Lin HF (2014) The analysis of bus commuters' travel characteristics using smart card data: the case of Shenzhen, China. In: 93rd annual meeting of the transportation research board
11. Chen J, Lv YK, Cui ML (2018) Estimating alighting stops of smart card public transportation passengers based on travel patterns. *Xi'an Univ Arch Tech (Natural Sci Edn)* 01(01): 23–29
12. Kieu LM, Bhaskara A, Chung E (2015) Passenger segmentation using smart card data. *IEEE Transp Intell Transp Syst* 16(3):1537–1548
13. Cheung F (2006) Implementation of nationwide public transport smart card in the Netherlands—cost-benefit analysis. *Transp Res Record: J Transp Res Board* 1971:127–132
14. Park JY, Kim DJ, Lim Y (2008) Use of smart card data to define public transit use in Seoul, South Korea. *Transp Res Rec* 2063(2063):3–9
15. Trepanier M, Tranchant N, Chapleau R (2007) Individual trip destination estimation in a transit smart card automated fare collection system. *J Intell Transp Syst* 11(1):1–14
16. Wang W, Attanucci JP, Wilson NHM (2011) Bus passenger origin-destination estimation and related analyses using automated data collection systems. *J Public Transp* 14(4):131–150
17. Ma XL, Wang YH, Chen F, Liu JF (2012) Transit smart card data mining for passenger origin information extraction. *J Zhejiang Univ Sci C Comput Electron* 13(10):750–760

18. Munizaga MA, Palma C (2012) Estimation of a disaggregate multimodal public transport origin-destination matrix from passive smartcard data from Santiago, Chile. *Transp Res Part C-Emerg Technol* 24:9–18
19. Sanchez-Martinez GE (2017) Inference of public transportation trip destinations by using fare transaction and vehicle location data dynamic programming approach. *Transp Res Record: J Tran Res Board* 2652(2652):1–7
20. Alsger AA, Mesbah M, Ferreira L, Safi H (2015) Use of smart card fare data to estimate public transport origin-destination matrix. *Transp Res Rec* 2535:88–96

Understanding Metro-to-Bus Transfers in Nanjing, China Using Smart Card Data



Miaoyan Zhang, Dongwei Liu, Yanjie Ji, and Yang Liu

Abstract Understanding metro-to-bus transfers are important in Chinese cities because metro systems in China have not formed extensive networks and metro riders rely heavily on bus services as a feeder and distributor to rail transit services. Given the wide application of automatic fare collection systems using smart card in Chinese transit systems, smart card data with boarding and alighting information has become a new source of data to understand transfer behaviors. This paper takes Nanjing, China as a case study and examines the possibility of using smart card data to understand specifically metro-to-bus transfer behaviors. Smart card data of a weekday and a weekend is used, with a total of over 6.5 million transaction records. The data analysis focused on identifying metro-to-bus transfers and how metro-to-bus transfer behavior (including transfer duration and time of the day) differ by card type, city region and weekend/weekday status. Results of the analysis show similarities and differences in metro-to-bus transfer behaviors across card holder groups, location of metro stations and between weekdays and weekend days. The transfer duration of the students and adults is shorter on weekdays than weekend days. Besides, the transfer duration in the suburbs is longer than urban areas and exurbs. Most metro-to-bus transfers on weekdays concentrate upon peak hours at 8:00 and 17:00. The reasons of these phenomena are discussed in the corresponding chapters. Limitations of the study and future directions are further discussed.

Keywords Metro-to-bus transfer · Smart card data · Time · City region · China

1 Introduction

Transfer constitutes an integral part of a public transit trip and heavily influences the attractiveness of public transit compared with other means of transportation [1]. It is also an important method for improving the accessibility of metro system which can reflect the difficulty for individuals to use urban public facilities, especially in

M. Zhang · D. Liu · Y. Ji (✉) · Y. Liu
School of Transportation, Southeast University, Nanjing 210096, China
e-mail: jiyanjie@seu.edu.cn

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_4

the construction of the metro system. China as a developing country is undertaking the most ambitious urban rail expansions to meet the growing travel demand that accompanies rapid urban growth. As of 2015, 26 cities in China have metros totalling 3,618 km, with a further increase under construction [2]. However, metros in many Chinese cities do not form extensive networks and a significant percentage of metro riders heavily rely on bus transportation to reach destinations [3]. In Nanjing, 33% of the metro riders used bus services to access destinations after their metro rides in 2013 [4]. With the extension of metro services, the number is likely to increase in 2015. In order to attract more travellers to use public transport, research should be conducted on the transfer behavior between metro and bus, which will have significant implications on individual trip-inducing and integrating demand management into the transportation planning.

Existing studies verified that transit users perceive transfer duration, especially waiting time during the transfer time significantly longer than they really are [5, 6]. These perceptions have negative implications for users' overall feelings about their mode [7, 8], and present a significant obstacle to increasing the competitiveness of public transit, which is much more environmentally friendly than the private automobile mode [6, 9, 10]. Given the billion-dollar scale metro investments in China, a decrease of the metro-bus transfer duration is important for improving the competitiveness of public transit and ensuring ridership successes of the metro projects. Although there were many studies on transfer time in the past, almost all of these studies use the questionnaire data, which cannot get the true transfer duration of passengers. Fortunately, Smart Card (SC) systems which were designed for fare collection can provide a detailed date record including card types and metro sites for many purposes [11]. It has the potential to replace many of traditional data collection (i.e., questionnaire survey) efforts and/or some elements of the planning process [12]. Nonetheless, relatively few studies were conducted using smart card data to analyze metro-bus transfer behaviors because only the off-boarding data is available [13]. Thus, the objective of this study is to better understand the distribution of passenger transfer time using smart card data. Generally speaking, different populations have different transfer duration in diverse regions and time, therefore this study will explore the distribution of transfer duration from three dimensions of users, city regions, and transfer time. Specifically, Smart card data for a week was collected from the Nanjing Citizen Card Company with over 26 million transaction records in total. Metro-bus transfers were identified using a predetermined maximum transfer time threshold and analyzed by type of the day (weekday and weekends), card type (student, adult, disabled, and elderly) and city regions. The analysis results have important policy implications for improving transit service operation, which in turn have implications for improving transit riders' travel experiences and increasing transit ridership.

In the following section, we review the relevant literature on using transit smart card data and metro-to-bus transfers. Then, the next section introduces the study area, the data and the identified method of transfer passengers. The 'Result and dissections' section presents the analyses of metro-to-bus transfer characteristics by users, site of metro stations and transfer time on weekday/weekend. Finally, the last

section summarizes the conclusions, policy implication, limitations, and suggestions for future research.

2 Literature Review

Researchers have investigated transfers between metro and other transportation modes using traditional data sources (i.e., questionnaire survey, GIS). Chen investigated the determinants of the demand for bicycle transfer at metro stations by conducting a survey of metro travelers' opinions towards bicycle transfer facilities at two metro stations of Nanjing [14]. Cherry used more than 300 surveys to assess passenger perceptions of metro-bus transfers in Bangkok, Thailand. They found that passengers expressed high levels of dissatisfaction with the overall transfer experience and that safety from crime and the distance between metro exits and bus stops have the greatest effect on passenger perceptions [15]. Navarrete and de Dios Ortúzar [7] investigated users' subjective perceptions of the metro-metro, metro-bus, and bus-bus transfer experiences and found that metro-bus transfers are valued more negatively by peak hour users than metro-metro and bus-bus transfers [7]. In addition, there is a general consensus that the quality and quantity of feeder/distributor bus services are important determinants of metro station ridership. Using GIS data, feeder/distributor bus lines were found to be associated with station ridership at the 0.01 significance level [16, 17] and the number of feeder/distributor bus lines has a reciprocal relationship with station boardings [18].

More recently, researchers have used smart card data for understanding various transit travel behaviors, including individual loyalty to transit services, transit trip distribution patterns, and route choice behavior. Chu and Chapleau [1] found that 80% of bus-bus linked trips have a transfer time of 18 min or less and the linked trips represent slightly above 10% of the total number of transactions in the network of Gatineau, Canada [9]. Trépanier et al. [19] used 5 years of smart card data of a medium-size transit authority in Gatineau, Canada and measured individual loyalty to the bus service (the length of the time an individual retained in the system). They found that individual loyalty is positively associated with residential density and the transit share in the area [19]. Ma et al. [20] used smart card data to detect transit riders' travel patterns and classify travel pattern regularities of transit riders in Beijing, China, and developed data mining algorithm are capable of processing massive smart card datasets within a tolerable elapsed time [20]. Sun and Xu [21] investigated analysis of travel time reliability and estimation of passenger route choice behavior based on Beijing metro smart card data [21]. Following that study, Sun and Schonfeld [22] developed a schedule-based passenger's route choice estimation model, taking fail-to-board phenomenon into consideration [22]. Zhu et al. [23] developed a method to calibrate urban rail transit assignment models using smart card data and a parallel genetic algorithm, which calibrates assignment model parameters by comparing observed and calculated travel time distributions [23].

Relatively few researchers used smart card data to study metro-bus transfer behaviors. Given that many metro and bus riders are using smart cards, smart card data showed great potential for describing the characteristics of public transit users including identifying transfer patterns, locating the critical transfer points that need improvement [12, 24]. Hofmann proposed iterative classification algorithm to identify transfer journeys based on waiting time information and spatial first/second boarding matrices with a large fleet (over 1000 buses) and data of 48 million magnetic strip card boardings [25, 26]. Shi and Hangfei [27] found that the bus-bus and metro-bus transfers are not many and most of the transfers were made within 30 min [27].

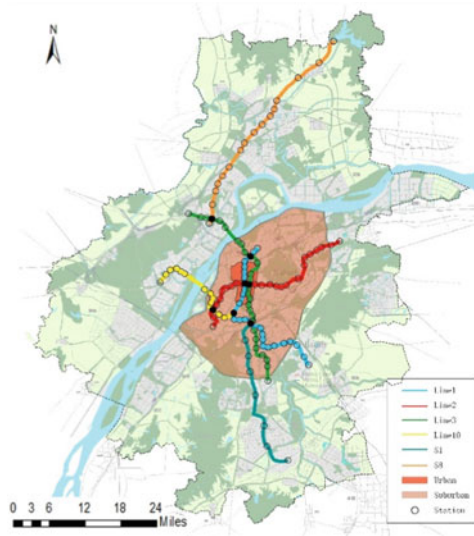
Previous research has made substantial progress in understanding transfer behaviors between metro and other transportation modes using traditional data sources. Substantial efforts have also been made towards the use of smart card data in understanding various transit travel behaviors. These studies come from different countries and regions. However, very few studies have specifically focused on the metro-to-bus transfer behavior in Chinese cities using smart card data. This study is a direct response to the knowledge gap to focus on metro-to-bus transfers specifically using smart card data from Nanjing, China.

3 Methodology

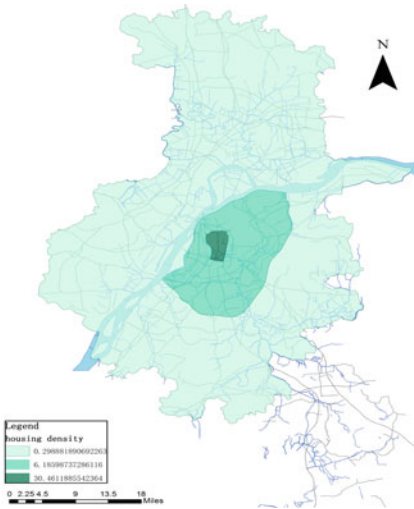
3.1 Study Area

Nanjing, the capital of Jiangsu province, is one China's most important commercial center and is adopted for this case study. Nanjing's population exceeds 8.2 million and the per capita gross domestic product is 127,264 CNY/year (nearly 18,927 US dollars/year) in 2016, with a total area of 6587 km² (11). The city is divided into three regions, namely urban, suburban, and exurban (13). Because the data of population and jobs are unable to obtain, the distribution of residence POIs and employment POIs was used to indirectly reflect the distribution of population and jobs respectively in these three areas. By counting the POIs in each area, we could get that 905 residential neighborhoods and 32,032 employment POIs are in urban area with an area of 29.7 km², 3388 residential neighborhoods and 74,871 employment POIs in suburban area with an area of 547.7 km², and 1248 residential neighborhoods and 52,399 employment POIs in exurban area with an area of 4165.5 km². In general, the urban area of the city could be characterized as a mixed land use with high densities of jobs and schools near rail transit facilities, while the suburban area is composed of single-function land use and has a low population density and poor rail transit facilities. It indicates that many commuters who work in the urban area are more likely to live in the suburban area.

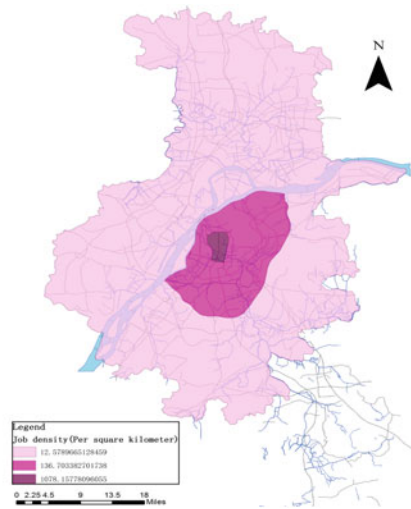
Figure 1a shows the distribution of metro stations and lines. A total of 113 metro stations are included in this study with 16 of them located in the urban region, 58



(a) Area Boundaries and the Existing Metro Lines



(b) Housing density



(c) Job density

Fig. 1 Study area in Nanjing, China

in the suburb area and 39 in the exurb area. Figure 1b shows the distribution of residential density in urban area, suburban area and exurban area. Figure 1c shows the distribution of employment density in the three areas. The darker the color is, the higher the density is.

3.2 Datasets

The smart cards collected from the Nanjing public transit system are similar to the SmarTrip card of the Washington metro, Ventra in Chicago, and Translink in the San Francisco Bay Area. The automatic fare collection (AFC) system can provide card type, the corresponding station ID and time stamps for each card ID when a passenger enters or alights the metro station. Smart-card holders were categorized as adults, the elderly, and students (less than 18 years old) and the disabled based on smart card type. The data used in this paper include a week of smart card transaction records from November 9, 2015 to November 15, 2015. The transaction data are recorded when passengers pass through the ticket gates or board the buses. Unfortunately, the alighting records of the bus are not provided so that the research is limited to a superficial aspect. The number of transactions in a typical weekday in Nanjing is about 3.0 to 3.5 million. For each transaction that occurs, several values are stored in the database. Table 1 shows a typical sequence of smart card transaction records generated by the system. For every smart card fare collection transaction, the date, time, card number, and card type are stored. In the case of a bus transaction, bus ID and line number are available. In the case of a metro transaction, ticket gate ID and metro station number are available, which provide locational information of transfers.

Since the departure time interval of metro and bus is likely to affect the transfer duration of passengers, this paper made a statistical analysis on the departure time interval of the representative bus lines around the metro stations in the three regions on weekdays and weekends, respectively. However, the bus route schedule provided by Nanjing public transit company is limited, only 7 bus lines are included in the urban area, 11 in suburban area and 5 in exurban area. The bus departure time intervals in different areas vary at different time period as shown in Table 2. It can be found that the time intervals in suburban areas are shorter than that in urban area on weekdays and weekends. And this time interval in urban and suburban at peak period is shorter than that at the reduced service period, followed by normal service period. On the other hand, we did statistics on the departure timetable of metro. The departure time

Table 1 A sequence of smart card transaction records (bus and metro)

Transaction date	Transaction time	Card type	Bus ID	Line number
2015-11-15	6:58:07	101	128142	106515100000
2015-11-15	18:29:14	101	141620	101502400000
2015-11-15	8:25:08	101	141618	101502400000
2015-11-15	9:04:02	101	148250	10151000000
Transaction date	Transaction time	Card type	Ticket gate ID	Metro station number
2015-11-15	08:22:13	101	20217363	0000051
2015-11-15	17:22:25	101	20069921	0000011
2015-11-15	17:22:25	101	20143362	0000030

Table 2 Departure time intervals of bus and metro (minutes)

Schedule	Bus						Metro	
	Weekday			Weekend			Weekday	Weekend
Area/time	Urban	Suburban	Exurban	Urban	Suburban	Exurban		
Peak period	5.7	4.7	11	8.3	3.2	11	4.3	4.6
Normal service period	9.0	6.7		11.2	7.3		7.4	6.3
Reduced service period	6.6	5.8		4.9	4.7		9.4	10

Note Peak period is 7:30 to 9:00 and 16:00 to 19:00; Normal service period is before 7:30, 9:00 to 16:00 and 19:00 to 20:30; Reduced service period is after 20:30

interval of the metro changes slightly between weekday and weekend. The results are shown in Table 2.

3.3 Identification of Transfer

As mentioned before, the data doesn't include the alighting records of the bus. Therefore, we focus on the metro-to-bus transfer analysis in order to explore its features in Nanjing. And transfer referred in this paper is restricted to the interchange from metro to bus (M-B transfer). It is very difficult to determine transfer behavior without supplemental information, such as GPS information and corresponding travel survey data. Seaborn (28) proposed a method to determine the transfer range based on the cumulative share curve of potential transfers within 60 min. In this way, the M-B transfer threshold is recognized between 15 and 25 min in London. Furthermore, according to prior research evidence and local statistics, most metro-to-bus transfers are within 30 min [4, 27, 28]. 60 min can be also used as a preliminary time limitation to analyze the M-B transfer in Nanjing. The previous survey shown that average egress time of metro involving the M-B transfer was within 20 min for commuters in Nanjing and the longest duration people could tolerate is 30 min [29]. Figure 2 illustrates the frequency and cumulative percentage distributions of potential M-B transfers which takes less than 60 min from metro alighting to bus boarding during a day. The potential transfers shorter than 25 min have reached 90%. In our paper, the potential transfer refers to the passengers who may have a transfer behavior but have not yet been identified. The cumulative curve increases sharply in the first 15 min and then the growth rate slows down. It reaches a low and stable rate at about 35 min. The transfers shorter than 15 min are considered to be pure M-B transfers, which means there is no other activity on the way from metro station to bus station. On the contrary, the potential transfers over 35 min are not pure transfers. Between 15

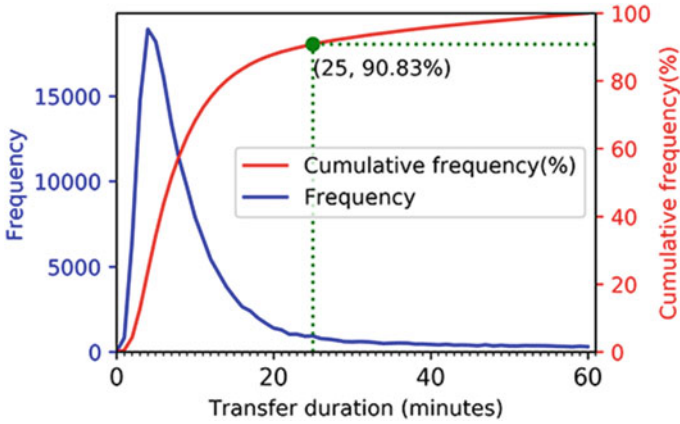


Fig. 2 The metro-bus transfer process defined in this research

and 35 min, there exist some incidental activity transfers. Buying a newspaper at the metro exit before bus boarding is such a non-travel activity that the two stages of travel can be linked. The time threshold varies from 15 to 35 min for different situations. Pure transfers and incidental transfers are both the interchanges that need to be discussed. After further analysis of cumulative curves in different metro stations, 25 min is adopted as an appropriate time threshold for M-B transfer in Nanjing as a whole, but every station may have a specific time threshold that ranges from 15 to 35. We identify all the M-B transfers each day of a week based on the transfer threshold.

4 Results and Discussions

4.1 Trip Patterns of Passengers

A week of smart card data gives the opportunity of measuring the daily trip patterns of smart card users. Following previous research [12], we code each card holder's daily transit trip patterns into series of alpha-numeric letters. The letter "B" means the bus mode and the letter "M" denotes the metro mode. The combination of TWO kinds of modes are used to represent the passengers' daily travel pattern. The former letter of the combination represents the first mode, and the latter letter represents the second mode. If the transfer relationship is established, no marks are between the two letters. However, if no transfer relationship is established, "&" is used to connect the two letters. Besides, if a passenger only used one mode to make a trip, a single letter is used to represent this. The number "3+" means equal or greater than 3 transactions in a day, for example, M&M (3+) means taking metro 3 times or over 3 times without any bus trip in a day. It is worth noting that each metro transaction may

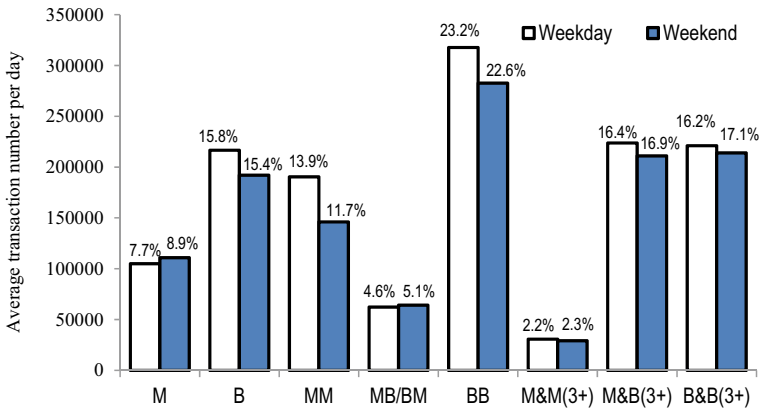


Fig. 3 Frequency and percentage distribution of the identified transit trip patterns

have internal metro-to-metro transfers that are not recorded in the smart card data. As shown in Fig. 3, we identified eight distinct daily transit trip patterns, represented as M, B, MM, BM/MB, BB, M&M (3+), M&B (3+), and B&B (3+).

The average percentage distribution of the identified transit trip patterns is shown as labels above the frequency columns in Fig. 3. As shown in the figure, the average number of smart card users in Nanjing on a weekday is about 1.367 million, of which 23.5% make a single trip (7.7% metro and 15.8% bus), 41.7% make two trips (13.9% metro only, 4.6% bus and metro, 23.2% bus only), and 34.8% make three or more trips (2.2% metro only, 16.4% metro and bus, and 16.2.0% bus only). About 30,718 people take metro 3 times or over 3 times on a weekday, which accounts for the lowest proportion of smart card users. The average number of smart card users in Nanjing on a weekend day is 1.250 million, which is slightly smaller than an average weekday. On the weekend day, single-trip card users are 24.3%, two-trip card users are 39.4%, and three-or-more-trip users are 36.3%. There are more bus-only card users on weekdays than weekend days. This is likely to be an indication of preference towards metro because people tend to have restricted destinations on weekdays and discretionary destinations on weekends. Potential metro-to-bus transfers were hidden in the BM/MB and M&B (3+) trip patterns who account for 21% of weekday users and 22% of weekend users.

4.2 Duration Analysis of Metro-to-Bus Transfers

Figure 4 shows the duration distributions of the identified metro-to-bus transfers, card type, location of metro stations and weekday/weekend status, among which, blue represents weekday and orange represents weekend. As stated previously, the smart cards were mainly divided into 4 types: Adult, Student (below high school),

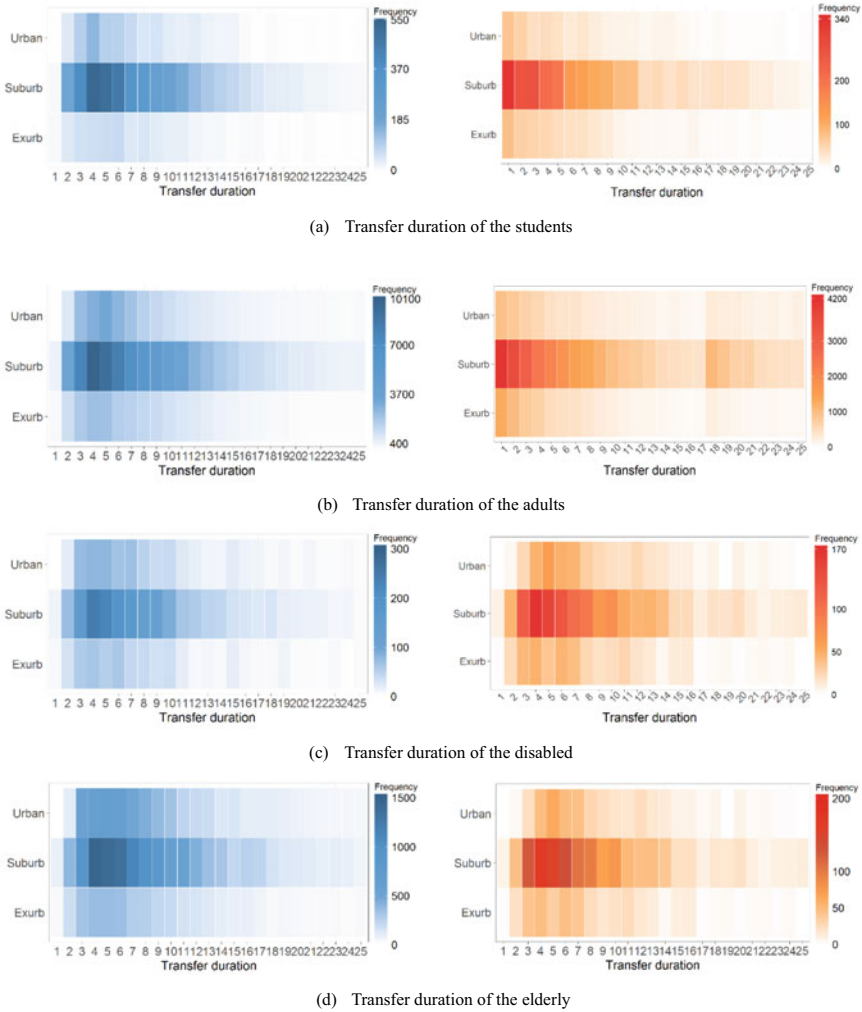


Fig. 4 Distribution of transfer durations of different card types in different city areas

Elderly (above 60 years old) and Disabled. The city of Nanjing was divided into 3 types: urban, suburb and exurb. We need to mention that whatever the color of figure is, the darker color always indicates the higher transfer volume. In Fig. 4, the transfer passenger flow in suburban area is much larger than that in urban and exurban area. It is understandable, because the density of the metro lines in suburban area is much lower than that in urban area, which means that passengers are less likely to arrive at their destinations with only one metro ride. At the same time, the transfer duration of passengers in suburban area are longer than that in urban and exurban area. The bus departure time in exurban area is earlier than that in urban district, which indicates that accessibility is the main factor affecting transfer duration. The walking distance

from the metro station to the bus station in exurban area is longer than that in urban area, which results in a long transfer duration. Besides, the transfer volume in exurban area is much higher than that in urban and exurban area. Thus, the bus congestion is likely to be another reason leading to a long transfer duration. For this reason, we further analyzed the transfer durations of different passenger groups on weekdays and weekends.

Figure 4a shows that the transfer duration of students on weekdays is significantly different from that on weekends. The transfer duration of students on weekend is reduced by about 2 min compared with that on weekday. Two reasons have led to this. On the one hand, the bus departure time in urban area on weekends is earlier than that on weekdays, which results in the decrease of the waiting time of students in urban area. On the other hand, because of the different travel purposes at the weekends, the transfer duration may be reduced by dispersal of students' departure time. As shown in Fig. 4b, the same rule can be found in the adult group. It indicates that students and adults show similar transfer rules under the constraints of their work and school schedule. In addition, the students in exurban area travel less, and the main reason is likely to be the less educational resources in exurban area.

Compared with these two groups, the elderly and the disabled groups show different transfer characteristics as shown in Fig. 4c, d. Their transfer duration has not decreased despite that the bus average departure time is shortened, which indicates that long walking distance is the main reason for their long transfer duration. We also found that the transfer volume of the elderly in urban area was relatively high, indicating that metro and bus are the main traffic modes used by the elderly.

4.3 Time Analysis of Metro-to-Bus Transfers

Figure 5 shows the distributions of the metro transaction occurring time (i.e., the moment when the fare was collected from smart cards in the metro trip segment that is the alighting time of the first trip stage by metro) by weekend/weekday status, city regions. As shown in Fig. 5, the proportion of metro riders who transfer from metro to bus is generally higher on weekdays than that on weekends. This may be attributed to the differences in the destinations and trip purposes between weekdays and weekends.

Figure 5a shows that the students' transfer time on weekdays is relatively concentrated. The morning peak is at 6:00–7:00, and the evening peak is from 15:00 to 18:00. The main reason is that the start of school hours is relatively fixed, while the end of school hours is different from primary school, junior high school and high school. Also, the student transfer volume in evening peak is higher than that in the morning peak, which shows that some students don't transfer in the morning peak. Since students who take metro to school usually have a long distance from home to school, it's almost impossible for them to walk or cycle, especially for the pupils. Therefore, we assume that those students who don't transfer in the morning are escorted to school by their parents using other traffic modes. According to the

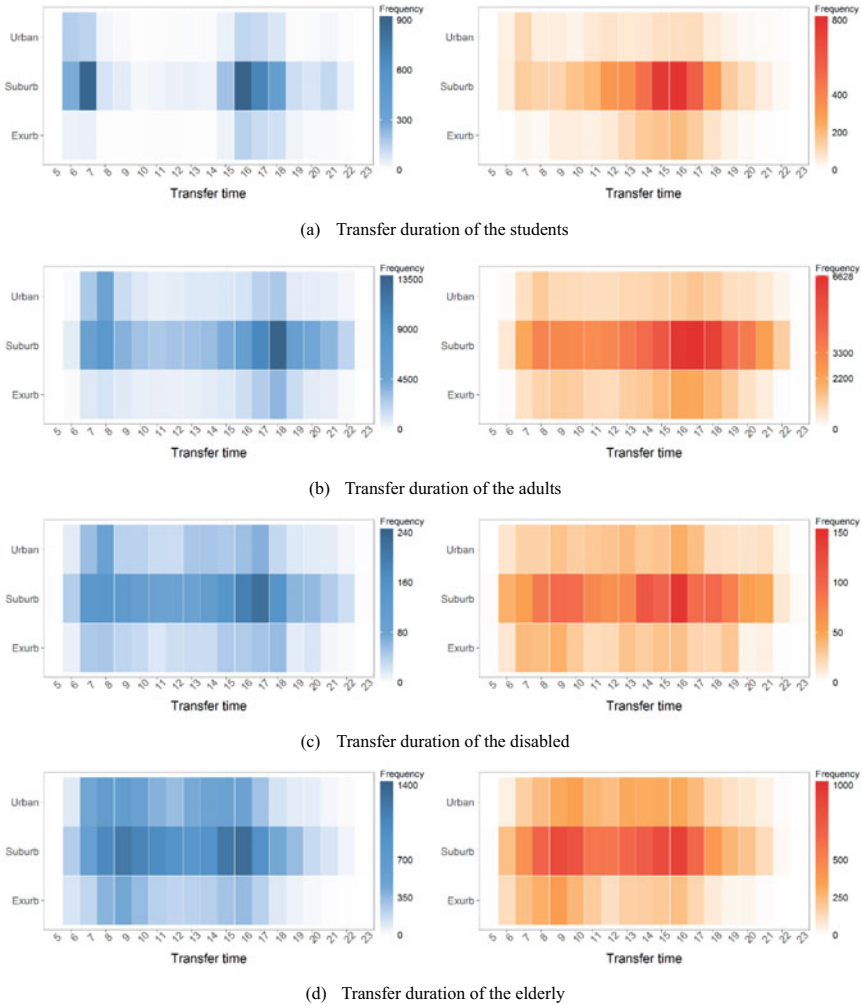


Fig. 5 Temporal metro trip distribution and transfer proportion in different city regions

2015 Nanjing resident travel survey, 57.8% of parents will escort their children to school, in which the ratio of car-use in long-distance travel is as high as 81.7%. It indicates that these students are very likely to be escorted to school by their parents in a car. In order to guide more students to take the metro to school, it is of vital importance to improve the transfer accessibility and shorten the transfer duration. It is noteworthy that the transfer peak of students is concentrated at 14:00–17:00 on weekday, which is different from that of other passenger groups.

For the adult group, the transfer morning peak is at 7:00–9:00, and the transfer evening peak is at 17:00–18:00, among which the transfer volume reaches the maximum value at 18:00. The distribution of adults' transfer time on weekends

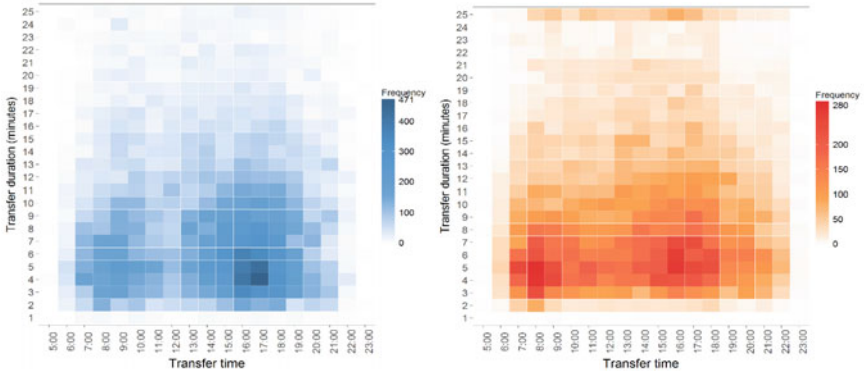
is similar than that on weekdays. The difference is that the transfer volume in urban and exurban area increases and the evening peak lasts 3 h from 15:00 to 18:00. It is possible because some adults still need to go to the weekend, and others travel for non-commuter purposes.

Different from the transfer time distribution of students and adults, the transfer time distribution of the disabled and elderly are more dispersed, and the transfer volume in the urban and exurban area is relatively increased. There is no obvious transfer morning peak on weekdays, and the evening peak is at 16:00–17:00, which is obviously earlier than that of adults. The transfer time distribution of the disabled on weekends is similar to that on weekdays. Different from the disabled, the elderly have obvious morning peak and evening peak, respectively at 8:00–10:00 and 15:00–16:00. And their evening peak is significantly earlier than that of the disabled and adults, the same as the students, which indicates that most elderly people may pick up children from school. In other words, those elderly people who have transfer behavior at the evening peak actually travel to pick up children to a great extent, which is consistent with China's national conditions. These discoveries are conducive to the arrangement of bus operation.

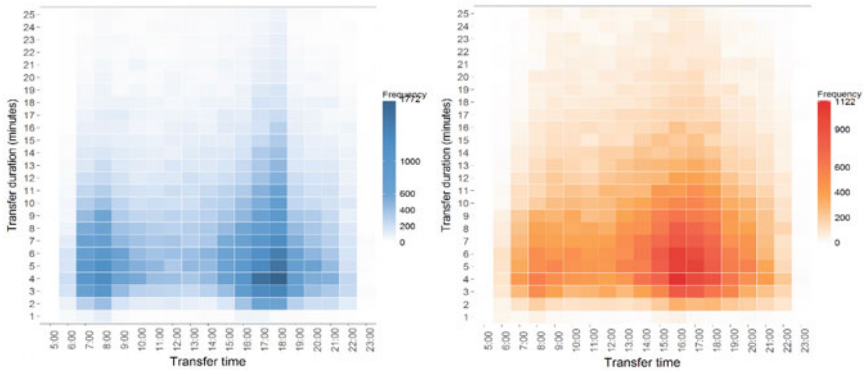
4.4 Time and Duration Analysis of Metro-to-Bus Transfers

We further investigated the distribution of passengers' transfer duration in different transfer time in these three areas. Figure 6 shows that the distributions of transfer duration in urban, suburban and exurban area are significantly different. Also, the distributions of transfer duration on weekdays and weekends in each area are also different. In Fig. 6a, the peak of transfer time on weekdays in urban area is at 16:00–17:00 with 4–6 min transfer duration, and most passengers can complete the transfer within 11 min. But on weekends, the peaks of transfer time in urban area are at 7:00–9:00 and 16:00–18:00. Because the passengers who transfer at morning peak are mainly the elderly and the disabled based on the analysis above, the result indicates that those passengers tend to make short-time transfer trips in urban area on weekend mornings. Besides, the transfer duration at transfer peaks on weekday increases significantly compared with it on weekends. We also found that the transfer time of some passengers is close to 25 min from 8:00 to 16:00. One possible reason is that the bus departure interval at off-peak time on weekends is 11 min. Another reason is that the walking speed of passengers decreases since they don't need to work at a fixed time.

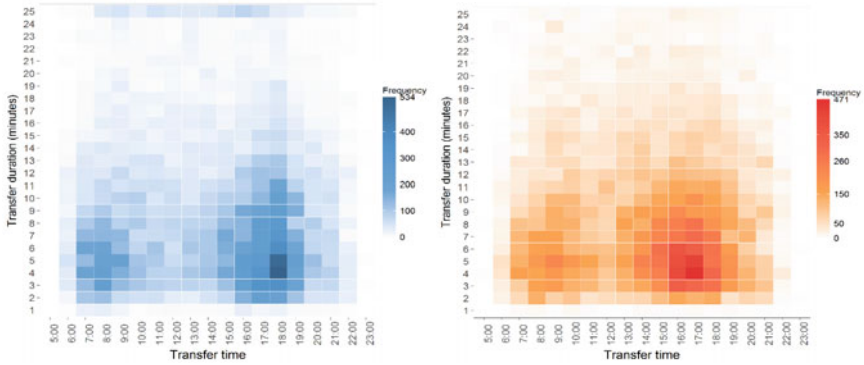
Figure 6b shows that the distribution of passengers' transfer duration in different transfer time in the suburban area is significantly different from that in urban area. On weekdays, there is an obvious morning peak and evening peak of transfer time. The morning peak is from 7:00–9:00 with transfer duration of 3–6 min. The evening peak is from 17:00–18:00 with transfer duration of 3–16 min. Excessive transfer volume is the main reason for the longer transfer duration during the evening peak because the average departure interval is 4.7 min, which is shorter than the departure



(a) Time and duration distribution in urban area



(b) Time and duration distribution in suburb area



(c) Time and duration distribution in exurb area

Fig. 6 The distribution of transfer time and duration in different city regions

interval in the off-peak period and low traffic period. Fewer passengers transfer in suburban area on weekend morning compared with weekdays. It is noteworthy that the transfer time on weekend afternoon lasts for a longer duration, from 14:00 to 18:00, which is obviously different from that on weekdays.

Figure 6c shows that the distribution of passengers' transfer time in exurban area on weekdays is close to that in suburban area. The transfer peak is at 17:00–18:00 with transfer duration of within 11 min. Because the bus departure interval is longer in this area, which means that most passengers in exurban area transfer according to the bus arrival time. Based on the previous analysis, most of these passengers are the elderly and the disabled. The passengers' transfer duration in the exurban area on weekends is close to that on weekdays, which may be caused by the same bus departure time interval. And the transfer peak in the exurban area on weekends is at 16:00–17:00.

5 Conclusions

The data used in this study is the smart card data of metro and bus in Nanjing, China during November 9–15, 2015. The analysis focused on identifying metro-to-bus transfers and how does metro-to-bus transfer behavior (including transfer duration and time of the day) differ depending on card type, site of metro stations and weekend/weekday status. Results from the analysis show similarities and differences in metro-to-bus transfer behaviors across cardholder groups and city regions between weekdays and weekend days. We found that because the job density and school density of urban area with a relatively low residential density are greater than suburb and exurb area, many residents live in suburb area and work or study in urban. For this reason, the number of passengers transferred in the suburb is much more than that in urban and exurb, which is the main reason leading to the long transfer duration in this area. In order to shorten the transfer duration in the suburb region, reducing the transfer volume of transfer time is an important way. On the one hand, it is possible to reduce cross-regional travel by improving the balance between occupation-housing and schooling-housing, such as implementing *the nearby enrollment policy*, but this change takes an extended period of time. On the other hand, policies can be formulated at the level of demand management to avoid passengers' concentrated time transfer. According to our conclusions, there are at least two ways that can reduce the transfer time of suburb passengers.

First of all, we conclude that the evening peak of suburban transfer is 17:00–18:00, and transfer duration lasts from 3 to 16 min. The main transfer passengers in this time period are adults and students, and their travel time depends mainly on the schedule of the work unit or school schedule. Previous research has pointed out that the tolerance of commuting time for most commuters is within 50 min in China, and the males tend to tolerate less commuting time compared with the females. For this reason, commuters, especially the male ones, are more likely to choose other travel modes rather than metro if there is a long transfer duration during the commuting

[30]. Therefore, it is necessary to propose a Stagger Shifts policy to alleviate transfer congestion. The transfer duration of passengers in the area decreased significantly after 18:00, work units can take flexible departure system. For example, take turns to make the get off work time of staff delay for about 10 min. In addition, a fare discount policy helps to guide passengers to travel stagger pea.

Secondly, the transfer time that passengers maintain a longer transfer duration (less than 10 min) in the suburb distribute a longer period of time, from 14:00 to 18:00. At present, Nanjing Bus Company's weekend departure time at early and evening peak is the same as the weekday. Although the average departure time interval in weekdays at 16:00–19:00 is only 3.2 min in suburb, but the departure interval is 7.3 min at 14:00–16:00, which will cause many transfer passengers to have to wait for the bus for a long time. In order to reduce the passengers' transfer duration, it is necessary to adjust the bus weekend departure time to shorten the departure time interval at 14:00–16:00 in suburb.

Furthermore, public bicycles are developing rapidly in China. Compared to private bikes, sharing bicycles allow users to use them on an as-needed basis, which eliminates any concern about maintenance costs and responsibilities [31]. Therefore, it is widely accepted by all kinds of people (young, adult, old people). For those transfer passengers who use public transit for a short distance, especially for young people and adults, it is a better choice to use public bicycles to transfer to the metro. Therefore, increasing the public bike around the metro station and increase bicycle delivery at 7:00–9:00 and 17:00–18:00 can reduce the transfer duration to some extent.

It is worth noting that besides the conclusion of reducing transfer duration in suburb areas, the paper also draws some conclusions which are beneficial to attract more residents to use public transport system. Firstly, we find that compared with the students and adults, the elderly have more transfer behavior in urban and exurb area, which shows that the metro and bus mode is the main way of the elderly people's daily travel. The addition of bus routes dedicated to the elderly helps to reduce the number and time of their transfer. Secondly, we found that the transfer volume of students in the weekday morning is far less than that in the afternoon, indicating that the students may be escorted by parents to school by other kinds of transport in the morning, especially by the car. Because the students' morning peak travel time is 6:00 to 7:00, the government can support to discount the public transit fare of student card before 7:00 based on student fare to encourage them to use the public transportation system. In addition, due to the elderly and the disabled tend transfer at a fixed time in the suburb area, the corresponding bus departure time can be provided according to their travel regularity.

This study relies upon the interval between metro and the subsequent bus transaction time to identify metro-to-bus transfers. The identified metro-to-bus transfer intervals do not include the time traveling from the metro platform to the exiting ticket gate, which could cause a significant underestimate of metro-to-bus transfer durations at large metro stations. Transfer behaviors of different card holders differ by city regions. Identifying factors underlying the differences are likely to be multifaceted, including tolerance to longer transfer duration across different groups of cardholders and city area, preferences of transfer behaviors, and destination types

between weekdays and weekend days across different card holders, etc. While identifying the factors is beyond the scope of this paper, the findings in the study point to the importance of future research on the subject. Dependent upon the underlying factors, the policy implications may be different. In addition, future research may consider using locational information of metro and bus transactions to identify metro-to-bus transfers more accurately.

Acknowledgment The financial support from the Projects of International Cooperation and Exchange of the National Natural Science Foundation of China (no. 51561135003), the Key Project of National Natural Science Foundation of China (no. 51338003), and the Project of National Natural Science Foundation of China (no. 51668029) is gratefully acknowledged.

References

1. Chu KKA, Chappleau R (2008) Enriching archived smart card transaction data for transit demand modeling. *Transp Res Rec J Transp Res Board* 2063(2063):63–72
2. Riedel H-U (2014) Chinese metro boom shows no sign of abating. *Int Railway J* 54(11)
3. Iseki H, Taylor BD (2009) Not all transfers are created equal: Towards a framework relating transfer connectivity to travel behaviour. *Transp Rev* 29(6):777–800
4. Nanjing Planning Bureau. Nanjing Transport Annual Report 2013
5. Dziekan K, Kottenhoff K (2007) Dynamic at-stop real-time information displays for public transport: effects on customers. *Transp Res Part A Policy Pract* 41(6):489–501
6. Watkins KE, Ferris B, Borning A, Rutherford GS, Layton D (2011) Where Is My Bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. *Transp Res Part A Policy Pract* 45(8):839–848
7. Navarrete FJ, de Dios OJ (2013) Subjective valuation of the transit transfer experience: the case of Santiago de Chile. *Transp Policy* 25:138–147
8. St-Louis E, Manaugh K, van Lierop D, El-Geneidy A (2014) The happy commuter: a comparison of commuter satisfaction across modes. *Transp Res F Traff Psychol Behav* 26:160–170
9. Alfred Chu K, Chappleau R (2008) Enriching archived smart card transaction data for transit demand modeling. *Transp Res Rec J Transp Res Board* 2063:63–72
10. El-Geneidy AM, Hourdos J, Horning J (2009) Bus transit service planning and operations in a competitive environment. *J Public Transp* 12(3):3
11. Pelletier M-P, Trépanier M, Morency C (2011) Smart card data use in public transit: a literature review. *Transp Res Part C Emerg Technol* 19(4):557–568
12. Jang W (2010) Travel time and transfer analysis using transit smart card data. *Transp Res Rec J Transp Res Board* 2144:142–149
13. Song Y, Fan Y, Li X, Ji Y (2017) Multidimensional visualization of transit smartcard data using space–time plots and data cubes. *Transportation* 3:1–23
14. Chen L, Pel A, Chen X, Sparing D, Hansen I (2012) Determinants of bicycle transfer demand at metro stations: analysis of stations in Nanjing, China. *Transp Res Rec J Transp Res Board* 2276:131–137
15. Cherry T, Townsend C (2012) Assessment of potential improvements to metro-bus transfers in Bangkok, Thailand. *Transp Res Rec J Transp Res Board* 2276:116–122
16. Zhao J, Deng W, Song Y, Zhu Y (2013) What influences Metro station ridership in China? Insights from Nanjing. *Cities* 35:114–124
17. Zhao J, Deng W, Song Y, Zhu Y (2014) Analysis of Metro ridership at station level and station-to-station level in Nanjing: an approach based on direct demand models. *Transportation* 41(1):133–155

18. Sohn K, Shim H (2010) Factors generating boardings at metro stations in the Seoul metropolitan area. *Cities* 27(5):358–368
19. Trépanier M, Habib KM, Morency C (2012) Are transit users loyal? Revelations from a hazard model based on smart card data. *Can J Civ Eng* 39(6):610–618
20. Ma X, Wu Y-J, Wang Y, Chen F, Liu J (2013) Mining smart card data for transit riders' travel patterns. *Transp Res Part C Emerg Technol* 36:1–12
21. Sun Y, Xu R (2012) Rail transit travel time reliability and estimation of passenger route choice behavior: analysis using automatic fare collection data. *Transp Res Rec J Transp Res Board* 2275:58–67
22. Sun Y, Schonfeld PM (2014) Schedule-based route choice estimation with automatic fare collection data for rail transit passengers. In: *Transportation research board 93rd annual meeting, 2014, vol 14-0834*
23. Zhu W, Xu R, Jiang Y-P (2014) Calibration of rail transit assignment model with automated fare collection data and a parallel genetic algorithm. *Transp Res Rec J Transp Res Board* 2442:20–28
24. Park J, Kim D-J, Lim Y (2008) Use of smart card data to define public transit use in Seoul, South Korea. *Transp Res Rec J Transp Res Board* 2063:3–9
25. Hofmann M, Wilson SP, White P (2009) Automated identification of linked trips at trip level using electronic fare collection data. In: *Transportation research board 88th annual meeting, 2009, vol 09-2417*
26. Hofmann M, O'Mahony M Transfer journey identification and analyses from electronic fare collection data. In: *Intelligent transportation systems, 2005. Proceedings. 2005 IEEE. IEEE, pp 34–39*
27. Shi X, Hangfei L (2014) The analysis of bus commuters' travel characteristics using smart card data: the case of Shenzhen, China. In: *Transportation research board 93rd annual meeting, vol 14-2571*
28. Seaborn C, Attanucci J, Wilson NH (2009) Using smart card fare payment data to analyze multi-modal public transport journeys in London
29. Yang M, Zhao J, Wang W, Liu Z, Li Z (2015) Metro commuters' satisfaction in multi-type access and egress transferring groups. *Transp Res Part D Transp Environ* 34:179–194
30. He M, Zhao S, He M (2016) Tolerance threshold of commuting time: evidence from Kunming, China. *J Transp Geogr* 57:1–7
31. Ji Y, Fan Y, Ermagun A, Cao X, Wang W, Das K (2017) Public bicycle as a feeder mode to rail transit in China: the role of gender, age, income, trip purpose, and bicycle theft experience. *Int J Sustain Transp* 11(4)

Research and Optimization of Vehicle Braking Efficiency Under Different Working Conditions Based on AMESim



Yiying Wei, Baoliang Ji, Ziyuan Xiong, Jingqiang Liang, Wei Liu,
and Zhen Huang

Abstract Brake efficiency is one of the most important factors that significantly affects driving safety. In this paper, a simulation model of the braking system was established by using AMESim. The accuracy of the simulation model was evaluated with the real vehicle tests. Based on the established model, the braking efficiency of the vehicles were analyzed under 15 different working conditions with different adhesion coefficient road surfaces at high speed, medium speed and low speed. The optimization of the parameters of the braking system was also proposed and studied. Research results show that the model is suitable for the study and optimization of brake efficiency under different working conditions. It lays a theoretical foundation for the research and optimization of automobile braking performance under the complex working conditions; shortens the development period of automobile brake system. At the same time, the model provides the theoretical support for the smart car adaptive cruise control (ACC) and automatic emergency braking (AEB) research.

Keywords Automobile engineering · Different operating conditions · Braking efficiency · AMESim

Y. Wei · B. Ji · Z. Xiong · W. Liu
School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan 430070, China

Y. Wei
Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan 430070, China

J. Liang
SAIC-GM-Wuling Automobile Co., Ltd., Liuzhou 545000, China

Z. Huang (✉)
School of Automation, Wuhan University of Technology, Wuhan 430070, China
e-mail: h-zhen@163.com

1 Introduction

With economic growth and technological development, the safety and comfort of vehicle attracts many researchers' attention. The vehicle braking efficiency is one of the important factors that significantly affects driving safety. Braking efficiency of vehicle refers to the ability of the vehicle to rapidly reduce the speed until it stops. The occurrence rate of traffic accidents will be greatly increased with the increase of braking distance of 0.1 m [1]. Therefore, good brake efficiency can reduce the probability of accidents, and improve the driving safety of vehicles [2].

The analysis of braking efficiency mainly utilizes real vehicle testing and simulation analysis. Bench test or road test is adopted for real vehicle test [3], and most of them are bench test [4, 5]. In terms of simulation, Wang et al. [6] established a vehicle model through MATLAB to analyze the braking efficiency and verified its correctness and stability successfully. Guo et al. [7] analyzed the braking efficiency of the ABS braking system based on the PID algorithm. The simulation results show that the PID controller can achieve control effectively in different road conditions. In addition, Xia et al. [8] analyzed the effect of the front tire slip rate on braking efficiency based on MATLAB. The experimental results show that setting a larger outer front wheel target slip can improve the vehicle braking efficiency of the vehicle, but its braking direction stability will be poor. Li et al. [9] analyzed the influence of tire wear on braking efficiency based on test and data fitting. Experiment results indicate that the degree of wear has a great effect on braking efficiency. Gerlich et al. [10] pointed out the influence of noise and temperature on braking efficiency during braking and proposed corresponding optimization methods. Milenkovi et al. [11] proposed a heat dissipation method after analyzing the influence of thermal conductivity of brake pad on braking efficiency. Kim et al. [12] modeled a certain vehicle based on MATLAB, and developed the braking system at the same time.

Human, vehicle and the environment are the three important factors that make up the driving of a vehicle. However, many scholars only simulate the relationship between the braking system itself and the vehicle braking efficiency, which ignores the influence of complex road status and the complicated working conditions on vehicle driving. Meanwhile, with the development of automatous driving technology, more people are concerned on the impact of road environment changes on vehicle safety. In this paper, a simulation model of the braking system was established by using AMESim. Based on the established model, the braking efficiency of the vehicles were analyzed under 15 different working conditions with different adhesion coefficients road surfaces at high speed, medium speed and low speed. Furthermore, the braking efficiency of vehicle is optimized after analyzing the influence of braking system parameters on braking efficiency. The research on the braking efficiency of five kinds of roads under high-speed condition is mainly based on the safe distance between cars on express way. The research under the medium speed condition is mainly based on the GB7258-2012 middle road test detection standard. The research under low speed condition is based on pedestrian-vehicle condition under E-NCAP and C-NCAP.

2 Method

The dynamic model of vehicle brake system consists of six parts: brake pedal, vacuum booster, hydraulic master cylinder, pipeline, brake and the vehicle itself. This paper only introduces the dynamic model of vacuum booster with two chambers, the vehicle dynamic model of brake wheel cylinder and the full vehicle dynamic model due to limited space of the paper.

2.1 Vacuum Booster Model

Figure 1 shows the vacuum booster that includes a vacuum chamber and an atmospheric chamber, and the model can be equivalent to a dual-piston pneumatic cylinder. Among them, the front pneumatic cylinder is equivalent to the vacuum chamber. The rear pneumatic cylinder is equivalent to the atmospheric chamber. The intermediate diaphragm is equivalent to the cylinder piston. The booster input push rod is equivalent to the atmospheric chamber push rod, and the booster output push rod is equivalent to push rod of vacuum chamber.

Dynamic analysis of the double-chamber:

$$F_0 = P_2(A_1 - A_2) - P_1(A_1 - A_3) \tag{1}$$

where F_0 is the output force of the intermediate diaphragm. P_1 is the working pressure of the vacuum chamber (MPa). P_2 is the working pressure of the atmospheric chamber (MPa). A_1 is the effective area of the diaphragm (mm^2). A_2 is the atmosphere chamber putter shank area (mm^2), and A_3 is the vacuum chamber putter shank area (mm^2).

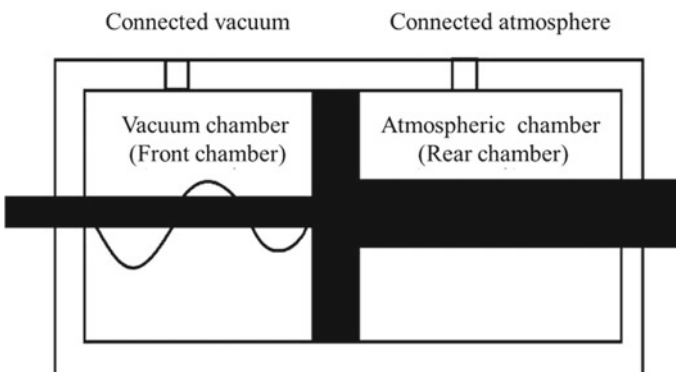
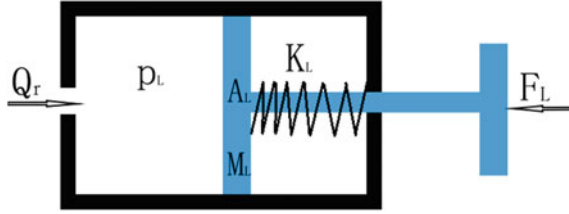


Fig. 1 Equivalent diagram of vacuum booster

Fig. 2 Dynamic model of brake wheel cylinder



2.2 Brake Wheel Cylinder Model

The brake wheel cylinder is a device that converts hydraulic energy into mechanical energy in the brake system. The floating caliper disc brake is used in this paper. And the dynamic model of brake wheel cylinder is shown in Fig. 2.

Without considering the roll-back effect of the piston contraction caused by the piston seal, the dynamic analysis of wheel cylinder is carried out:

$$M_L \ddot{X}_L = P_L A_L - C \dot{X}_L - K_L X_L - F_L \quad (2)$$

The liquid compression formula:

$$\dot{P}_L = E \frac{Q_r - \dot{X}_L A_L}{V_L + X_L A_L} \quad (3)$$

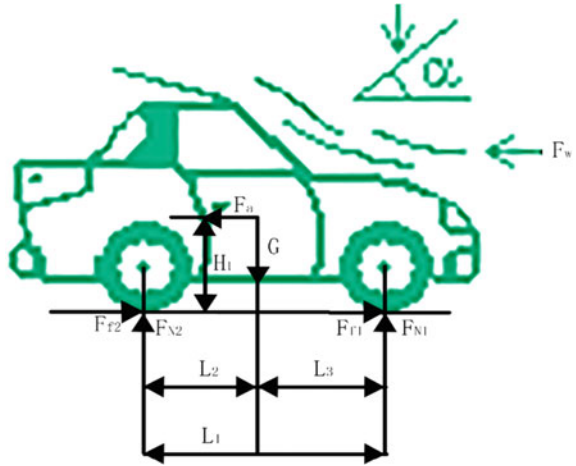
where C is the piston damping coefficient ($N/(mm/s)$); M_L is the mass of the brake pad (kg); X_L is the displacement of the brake pad (mm); P_L is the brake pressure in the cylinder chamber of the brake wheel (MPa); A_L is the cross-sectional area of brake wheel cylinder piston (mm^2); K_L is the equivalent spring stiffness of the rubber seal ring of the brake wheel cylinder (N/mm); F_L is the contact force between the brake pad and the brake disc (N); E is the bulk modulus of elasticity of brake fluid (MPa); Q_r is the inflow flow of brake wheel cylinder (mm^3/s); V_L is the initial volume of brake wheel cylinder (mm^3) [13, 14].

2.3 Vehicle Model

Figure 3 shows the forces of the vehicle dynamic model include air resistance— F_w , ground braking forces— F_{f1} , F_{f2} , ground normal reaction forces— F_{N1} , F_{N2} , rolling resistance- f and slope resistance— F_0 . This paper just analyzes the braking of the vehicle on the horizontal straight road, so the ground slope is 0, and the slope resistance is 0.

The dynamic equation of automobile braking could be written as:

Fig. 3 Dynamic model of vehicle braking



$$F_a = -F_{f1} - F_{f2} - F_w - f \tag{4}$$

$$F_{f1} = \varphi \frac{G}{L_1} (L_2 + \varepsilon H_1) \tag{5}$$

$$F_{f2} = \varphi \frac{G}{L_1} (L_3 + \varepsilon H_1) \tag{6}$$

$$F_w = \frac{1}{2} C_D A \rho V_r^2 \tag{7}$$

$$f = mg\mu \tag{8}$$

$$F_a = ma \tag{9}$$

$$G = mg \tag{10}$$

where F_a is the external force on automobile (N); F_{f1} , F_{f2} is the braking force of the front and rear wheels of the automobile (N); F_w is the air resistance (N); f is the rolling resistance (N); φ is the road adhesion coefficient; G is the gravity (N); L_1 is the wheel base (m); L_3 and L_2 are the distance between the center of gravity of the vehicle and the front and rear axles of the vehicle (m); ε is the severity of braking; H_1 is the height of the center of gravity; C_D is the air resistance coefficient; A is the automotive frontal area (m^2); ρ is the air density (kg/m^3); V_r is the relative speed of vehicle and air (m/s); m is the mass of the vehicle (kg); g is the acceleration of gravity (m/s^2); μ is the rolling resistance coefficient; a is the acceleration of vehicle (m/s^2).

2.4 Simulation of the Braking Model

Figure 4 shows the model established in the software of AMESim, and the model is used to analyze the braking system and its braking efficiency. A real vehicle test according to the parameters of a certain brand of SUV vehicle was conducted to verify the correctness of the model and simulation results. Only the parameters of brake and vehicle are listed due to limited space of the paper, as shown in Table 1.

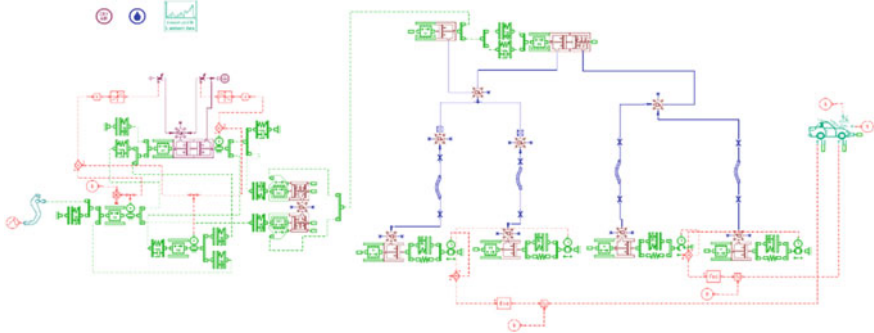


Fig. 4 AMESim model of vehicle brake system

Table 1 Main parameters of brake and vehicle

Mass of front/rear brake pads (kg)	0.7/0.1945
Diameter of front/rear brake wheel cylinder (mm)	54/32
Effective radius of front/rear brake wheel cylinder (mm)	122/119
Equivalent stiffness of front/rear brake piston seal (N/mm)	4/4
Front/rear brake clearance (mm)	0.15 / 0.15
Equivalent spring stiffness of contact model (N/m)	Front: 6×10^7 /rear: 3×10^7
Deformation when the damping reaches the maximum (mm)	0.15
Equivalent damping maximum damping coefficient (Ns/m)	Front: 6×10^5 /rear: 3×10^5
Vehicle weight (kg)	No load: 1431/full load: 1905
Front axle load (kg)	788
Rear axle load (kg)	643
Tire specifications	215/60 R17
Wheelbase (mm)	2750

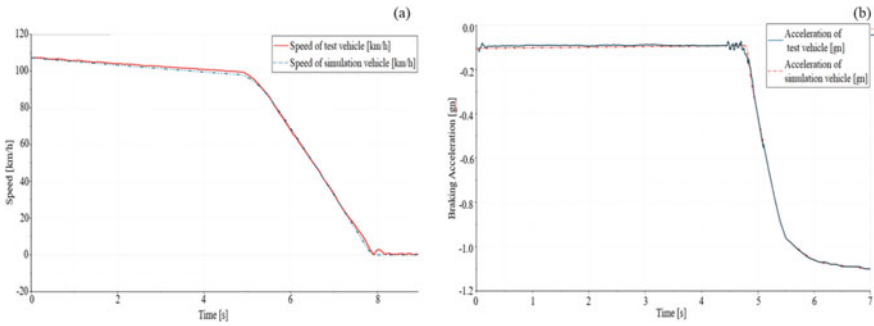


Fig. 5 Comparison of acceleration and speed between test vehicle and simulation vehicle

Figure 5 shows the acceleration and speed comparison between the test vehicle and the simulation vehicle. It can be concluded that there is a slight fluctuation at the beginning of the vehicle braking, and the simulation results have high coherence with the test results. The main reason for the fluctuation is the body limitation. When the driver starts braking, the vehicle will shake back and forth, resulting in a slight fluctuation of the acceleration of the test vehicle. The curve of the test results does not completely coincide with the simulation result, but the error is within the allowable error range of the test. Therefore, the rationality and accuracy of the simulation model have been proved.

3 Results and Discussion

When the vehicle is moving, the three factors of human, vehicle and environment will cause the multiplicity and complexity of driving conditions. The research objects of this paper are mainly 15 typical working conditions composed of three speeds and five road surfaces. Three vehicle speeds are high-speed (80 km/h), medium speed (50 km/h) and low speed (20 km/h). Five kinds of road surfaces are as follows: (1) asphalt road with adhesion coefficient-1.0; (2) wet asphalt road with adhesion coefficient-0.8; (3) gravel road with adhesion coefficient-0.68; (4) wet gravel road with adhesion coefficient-0.5 and (5) snow road with adhesion coefficient-0.2 [15]. The elimination time of vehicle brake clearance is 0.4 s, the rise time of vehicle brake force is 1 s, and the maximum pedal brake force is 500 N [16].

3.1 Different Road Conditions at High Speed

Vehicle driving at high speed (80 km/h) are usually on highways or wider roads. On this kind of road, the driver has a narrow line of sight and cannot handle the driving

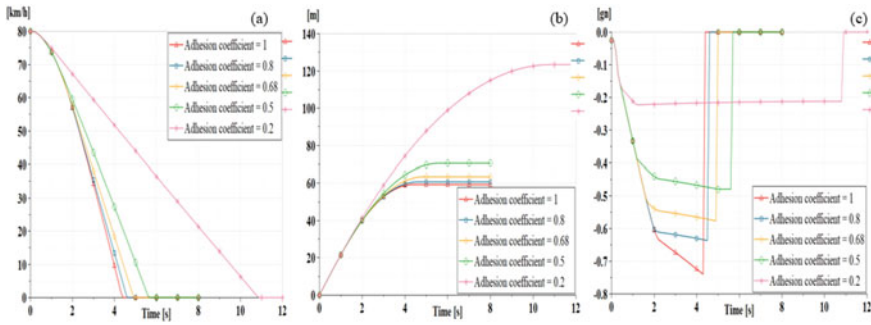


Fig. 6 Different braking speeds, braking distances, and braking accelerations at high-speed (80 km/h). On asphalt, wet asphalt, gravel and wet gravel and snow road, the braking distance of vehicles are 46, 50, 53, 60 and 120 m, and the braking time of vehicles are 3.5, 3.9, 4.2, 4.9 and 10.5 s

conditions of the surrounding vehicles in a timely and accurate manner. Accidents are prone to occur [17]. Therefore, the requirements for braking efficiency is relatively high. The simulation results of five kinds of road surfaces under the high speed (80 km/h) of vehicle braking are shown in Fig. 6. The curves in the figure are braking speed, braking distance and braking acceleration, respectively.

According to the research, the average reaction time of drivers is 1.5 s [18]. Therefore, the total braking distance of the vehicle from the driver’s intention to stop at high speed is 79 m, 83 m, 86 m, 93 m and 153 m, respectively. Among them, the braking distance of the vehicle on the asphalt road, wet asphalt road, gravel and wet gravel road is less than the general safe distance of 100 m on the highway, which shows that the vehicle can perform a better braking effect in the general conditions of high-speed conditions. However, the total braking distance reaches 153 m on the snow road, which means that it is necessary to maintain a straight distance of 153 m or more with the preceding vehicle or other obstacles to avoid collision when driving on the snow road at high speed.

3.2 Different Road Conditions at Medium Speed

Medium speed (50 km/h) is one of the most common vehicle speeds and the running speed of vehicle in most cities. The braking efficiency can be applied to most driving conditions. In the analysis of braking efficiency of passenger vehicles, the initial speed specified by the national standard is 50 km/h, and the detailed data is shown in Table 2 [19]. The calculation formula of full average deceleration is [20]:

$$MFDD = \frac{V_b^2 - V_e^2}{25.92 \times (s_e - s_b)} \tag{11}$$

Table 2 Emergency braking performance requirements

Motor vehicle type	Initial brake speed (km/h)	Braking distance (m)	MFDD (m/s^2)	Adhesion coefficient
Passenger vehicle	50	≤ 38.0	≥ 2.9	≥ 0.7

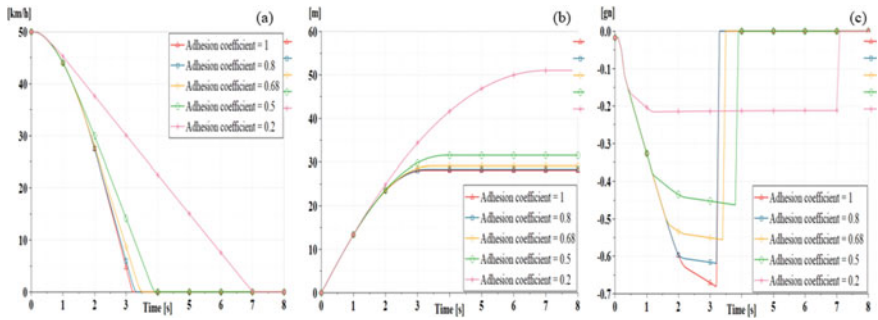


Fig. 7 Different braking speeds, braking distances, and braking accelerations at medium speed (50 km/h). The braking distance of the vehicle on asphalt, wet asphalt, gravel and wet gravel and snow road are 21 m, 22 m, 24 m, 27 m and 50 m respectively, and the calculated average deceleration of full emission are $7.7 m/s^2$, $7.6 m/s^2$, $6.2 m/s^2$, $6.0 m/s^2$ and $2.1 m/s^2$

where MFDD is the average fully deceleration (m/s^2); V_b is the test vehicle velocity, which is equal to $0.8V_0$ (km/h); V_e is the test vehicle speed, which is equal to $0.1V_0$ (km/h); V_0 is the initial braking speed of the test vehicle (km/h); S_b is the distance (m) when the test vehicle speed is from V_0 to V_b ; S_e is the distance (m) when the test vehicle speed is from V_0 to V_e .

The simulation results of five kinds of road surface under the medium speed (50 km/h) of vehicle braking are shown in Fig. 7. The curves (a, b, c) in the figure are braking speed, braking distance and braking acceleration respectively.

The data shows that the test vehicle has good average deceleration and short braking distance when driving on asphalt, wet asphalt, gravel and wet gravel roads at medium speed, which meets the testing standards of GB7258-2012 mid-road test. The test vehicle can brake faster and better and avoid collisions effectively when driving on asphalt, wet asphalt, gravel and wet gravel roads at medium speed. On other hand, the average deceleration of the test vehicle on the snowy road is small and the braking distance is large, which will lead to the vehicle collision easily under this working condition.

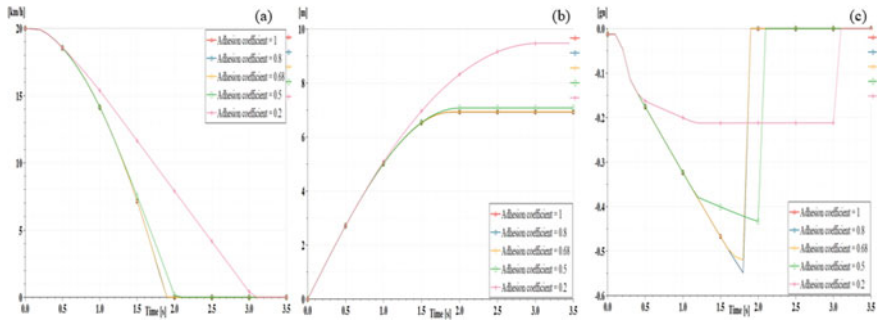


Fig. 8 Different braking speeds, braking distances, and braking accelerations at low-speed (20 km/h). The braking distance of the vehicle on the on asphalt, wet asphalt, gravel and wet gravel and snow road are 5.4, 5.6, 5.7, 6.1 and 9.6 m, the braking time are 1.4, 1.5, 1.6, 1.7 and 3.0 s

3.3 Different Road Conditions at Low-speed

Low-speed vehicle driving (20 km/h) conditions mainly occur when road traffic is congested or the vehicle is driven on rugged mountain roads. At this time, the distance between vehicles is short or there are many bends and obstacles. Therefore, there is a great demand for the braking efficiency of vehicle, which requires the vehicle to stop in a short distance. Figure 8 shows the simulation results of different road conditions at low-speed (20 km/h). The curves in the figure are braking speed, braking distance and braking acceleration respectively.

According to the data, the braking distance of the vehicle running under five kinds of roads at low speed is not more than 10 m, but the braking distance of the snow road is still very large. This shows that the driver needs more safety distance to avoid collision than usual when driving on the snow. Through the application of five working conditions data under 20 km/h in the test of AEB, the following conclusions can be drawn: When the vehicle brakes under the AEB Pedestrian working condition in E-NCAP [21], the vehicle can avoid collision with pedestrians on asphalt, wet asphalt, sand and wet sand; On snowy road, vehicles can avoid collisions under CVFA-50, and the braking distance of the vehicle should be more than 9.6 m away from the pedestrian’s horizontal driving distance under other working conditions; while in pedestrian test of C-NCAP, vehicles can avoid traffic very well on these five roads [22].

3.4 Parameter Adjustment and Efficiency Optimization of Brake System

Based on the above simulation data, the braking efficiency of the vehicle is better when running on asphalt, wet asphalt, gravel and wet gravel roads, and poor when

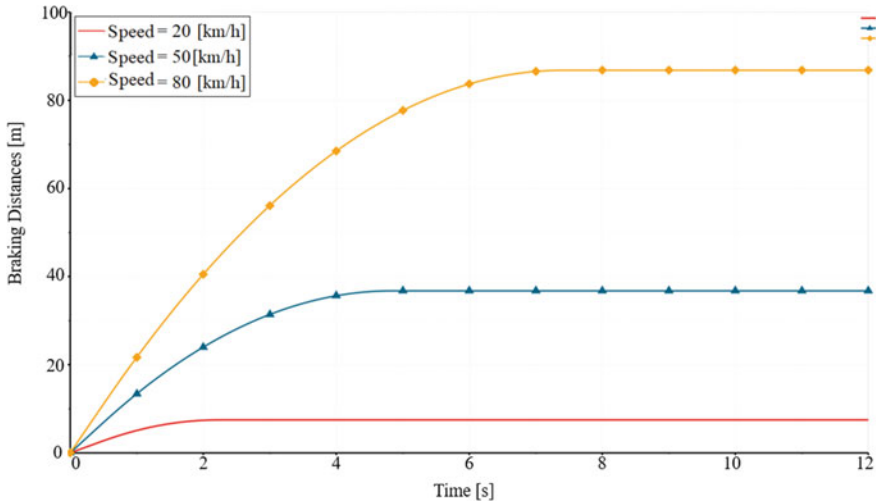


Fig. 9 Optimized braking distance of vehicle on snow road at different speeds

driving on snow. Therefore, the optimization of braking efficiency is mainly aimed at three different working conditions under snow road.

Analysis of simulation data shows that the main reason for the larger braking distance is the smaller coefficient adhesion on snow road. Therefore, the vehicle tires are changed to tires with a smaller effective radius and a greater coefficient of friction with the road surface, while increasing the equivalent stiffness of the brake caliper body deformation, so as to improve the braking efficiency of the vehicle. After three parameter optimizations, the braking distance on the snow road is 87 m at high speed, 37 m at medium speed and 7.5 m at low speed, as is shown in Fig. 9.

The optimization results show that: (1) under high-speed conditions, the total braking distance of the test vehicle is 120 m, which is 33 m shorter than the unoptimized; (2) under medium-speed conditions, the vehicle’s braking distance is 13 m shorter than the unoptimized. And, calculated the average deceleration of the vehicle at the full emission site is 3.2 m/s; (3) under low speed conditions, the braking distance is 2.1 m shorter than the unoptimized. According to the experimental results, the braking efficiency of the test vehicle has been improved to about 128% of the original. And it also meets the requirements of braking distance and full average deceleration in the GB7258-2012 road-test standard under medium-speed snow conditions.

4 Conclusion

Based on the whole vehicle model, this paper researches the braking efficiency of the vehicles under 15 different working conditions with different adhesion coefficients road surface at high speed, medium speed and low speed. Optimized the parameter

adjustment for the poor braking efficiency in snow conditions, which increased the braking efficiency by 128%. The research results show that the model can be used to analyze the braking efficiency under different working conditions and make optimization alter, which lays a research foundation for the research of braking efficiency of vehicles under more complex working conditions.

Meanwhile, this paper combines the experiment and simulation to ensure the accuracy of simulation results, which could reduce the research cost, shorten the research period and provide a research basis for analyzing different braking conditions of different vehicles. In addition, the simulation model can also be used to analyze the feeling of brake pedal under different conditions of traditional brake system.

Prospect: The next step is the vehicle drive system and complex road models can be add to the model, thus simulate the driving process of the vehicle on the complex road for a long time and a long distance. In addition, the vacuum booster can be replaced with an electronic booster. The analysis of braking intention can be added to realize the analysis of braking condition under unmanned vehicle driving.

Acknowledgements This work was supported by the National Key Research and Development Plan (No. 2017YFB0102500) and the National Key Research and Development Plan (No. 2019YFB1600803).

References

1. Li WT, Tian Z (2007) Programmed design and calculation of vehicle braking efficiency. *Mod Manuf Technol Equip* 3:38–38
2. Zhao YN (2018) Simulation analysis of braking performance of vehicle at different speeds. *Internal Combustion Engine Parts* 06:64–66
3. Xu WB, Zang MY, Yu SH (2014) Research on the test standard of passenger car braking performance. *Automob Technol* 8:38–41
4. Shi MM (2016) Comparison of test methods for automobile braking performance. *Auto Maintenance Repair* 1:80–81
5. Xu YB (2017) Discussion on the test of automobile brake performance. *China Plant Eng* 10:59–60
6. Wang P, Jiang T (2016) Modeling and simulation of automobile brake characteristic based on MATLAB. *Mech Eng* 2:66–68
7. Guo XD, Guo WL (2016) Simulation analysis of automobile braking performance based on PID control. *Transp Energy Conserv Environ Prot* 12(1):30–36
8. Xia CG, Ren YW, Chen S (2017) An analysis on braking efficiency and direction stability based on vehicle front wheel slip rate. *Mach Des Manuf* 12:31–34, 39 (2017)
9. Li J, Wei L, Zhang WH et al (2013) Tires wear degree influence on vehicle braking efficiency. *J Chongqing Jiaotong Univ (Nat Sci)* 2:335–338
10. Juraj G, Mykola G, Kateryna K, Olga P, Tomas L (2010) Noise and temperature reduction in the contact of tribological elements during braking. *MATEC Web Conf* 2018:157
11. Milenkovi PD, Jovanovi SJ, Jankovi AS et al (2010) The influence of brake pads thermal conductivity on passenger car brake system efficiency. *Therm Sci* 14:221–230
12. Kim S, Kim J, Sung G et al (2015) Evaluation and development of improved braking model for a motor-assisted vehicle using MATLAB/simulink. *J Mech Sci Technol* 29(7):2747–2754

13. Gua H, Hao WT, Zhan J (2013) A vacuum booster model for brake pedal feeling analysis. *Adv Mater Res* 622–623:1248–1252
14. Meng DJ, Zhang LJ, Fang MX et al (2014) Vacuum booster-brake master cylinder system dynamic model for brake pedal feel. *J Tongji Univ (Nat Sci)* 42(12):1897–1903
15. Breuer B, Bill K et al (2011) *Brake technology handbook*. China Machine Press
16. Zhang Y (2017) *Automobile performance and evaluation*. Chemical Industry Press, Beijing, p 01
17. Liu D, Ma SQ, Niu XJ (2008) Analysis on the characteristics of highway traffic accidents in China. *J Chinese People' s Publ Secur Univ (Sci Technol)* 14(4):65–68
18. Lin L, Zhu XC, Ma ZX (2014) Driver brake reaction time under real traffic risk scenarios. *Autom Eng* 36(10):1225–1229, 1253
19. Zhang AH (2018) *Automobile performance testing technology*. Beijing University of Technology Press, Beijing, p 08
20. Zhang SQ (1998) Mean fully developed braking deceleration MFDD. *Heavy Truck* 49(6):22–25
21. Gao JD, Zhang H, Gao BL et al (2017) Test protocols of autonomous emergency braking systems. *Auto Eng* 1:11–15
22. Wu J, Xiang GL, Yang JH (2018) Development and measurements of automotive autonomous emergency brake (AEB) pedestrian testing system. *J Autom Safety Energy* 9(4):401–409

Effects of the Layout of Exhaust Port on the Performance of Smoke Control for Ultra-Wide Immersed Tube Tunnel



Ying Li, Shenglin Mu, Xiaoyang Ni, Zhongchen Niu, and Kaixuan Tang

Abstract Smoke control is a crucial issue in tunnel fire prevention and control, especially in the ultra-wide tunnels, which is arousing more and more attentions. In this paper, the performance of three different exhaust port setting methods on the fire smoke control in the four-lane tunnel of immersed tube is compared and analyzed based on CFD numerical simulation. In the super wide immersed tunnel, increasing the number of smoke outlets on the transverse side of the tunnel can effectively reduce the thickness of the smoke layer and the maximum temperature of the vault under the top centralized smoke exhaust strategy. The efficiency of smoke exhaust of the outlet along the tunnel obviously decreased with its distance far away from the fire. In the design of smoke prevention and exhaust for super wide section immersed tube tunnel, we can consider reducing the gap between the exhaust fume and increasing the number of horizontal exhaust fume to improve the effect of smoke control in tunnel.

Keywords Immersed tunnel · Exhaust port · Smoke layer thickness · Efficiency of exhaust smoke

1 Introduction

With the solution to technical issues and its improvement, there is more and more attention to the immersed tube tunnel, which gradually becomes the preferred method of construction. Nevertheless, fire is the greatest risk in underwater tunnels while

Y. Li
Naval University of Engineering, Wuhan 430074, China

S. Mu · Z. Niu · K. Tang (✉)
School of Safety Science and Emergency Management, Wuhan University of Technology, Wuhan 430070, China
e-mail: kaixuant@whut.edu.cn

X. Ni
Faculty of Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China
e-mail: xy_ni@163.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_6

hyperthermia, poisonous and harmful smoke caused by fire is the key factor of causing deaths. As the section width of the immersed tube tunnel is increasing, if it takes more time for smoke to spread longitudinally along the tunnel, the smoke-spreading range will also be enlarged. The above-mentioned issues bring new challenges into smoke control in tunnels [1–4].

So far, designs for smoke exhausting during the fire in tunnels have been largely studied [5–8]. Given the complexity and high cost of field experiments in tunnels, completed researches rely mostly on numerical simulation. Wang et al. [9] used numerical simulation analysis to study the influence of different numbers of exhaust port both upstream and downstream on smoke in the tunnel. Results show that the higher the number of smoke outlets are, the higher the maximum temperature at the fire source is. Along with the increasing numbers of exhaust ports at the upstream of fire source, the height of smoke layer at the downstream area will increase accordingly. Liu et al. [10] proposed a simulation research on the critical wind speed based on centralized smoke extraction system. The research simulates different fire location and critical wind speed at different exhaust ports. It's indicated that the modified gradient value is of great importance when adopting either sidewall or top smoke exhaust strategy. Xu et al. [11] did a research on reasonable smoke extraction value of smoke-exhausting system in tunnels by using numerical simulation and experiment method. The results find a more reasonable smoke extraction volume of 140 m³/s. Wang et al. [12] analyzed characteristics and harm of fire smoke in the undersea tunnel combing actual condition of Jiaozhou Bay subsea tunnel in Tsingtao.

It's observed that enormous researches are almost on the influence of the setting location of exhaust ports in the tunnel, opening area and spacing on effect of smoke extraction [13–15]. There has been relatively little research on the way of setting up exhaust port for tunnels. For a four-lane immersed tube tunnel with a wider section, it's hard to achieve the expected smoke-exhausting effect by traditional ways of setting exhaust ports either at sidewall of at the top [16, 17]. In this paper, the focus is on the influence of different setting ways of exhaust ports on the effect of smoke extraction when adopting the top smoke exhaust strategy in the four-lane immersed tube tunnel. Giving parameters such as smoke layer, vault temperature and heat extraction efficiency, this paper introduces comparative analysis on the effect of different setting ways of exhaust ports on smoke control, which provide reference for design schemes of smoke extraction system in the immersed tube tunnel.

2 Model Construction

The research method of smoke control within fire period focus mostly on physical model test and computer numerical simulation. Numerical simulation is of many advantages such as great flexibility, short period and low cost [18–22]. This paper is based on the fire dynamics simulation software FDS (Fire Dynamics Simulator) that is developed to simulate the fire scenario by US NIST [23]. FDS is mainly used to

analyze smoke flow during fire period and the solution to heat transmission, which is widely accepted in the construction and fire-fighting domain.

When smoke is flowing in the tunnel, the wider the section of tunnel is, the longer the time needed for high-temperature smoke to spread longitudinally is. Accordingly, the single-row exhaust port setting longitudinally has less control over high-temperature smoke around the tunnel sidewall, so that effect of smoke exhaust will become worse along with time. As for the underwater tunnel with an ultra-wide section, to transversely tighten the lateral control over smoke, it's practicable to do focused smoke exhaust by longitudinally setting two or more rows on the top of the tunnel.

In the case of constant smoke extraction volume, three cases of exhaust ports are set along the tunnel as follow: in one row (Case 1), in two rows (Case 2) and in three rows (Case 3). At the same time, it should also respectively correspond to two cases of setting exhaust ports: on one side or on both sides of the fire source.

When set in Case 1, the first exhaust port is also the closest to the downstream fire source, the distance of which to the fire source is 10 m. As is shown in Fig. 1, when exhaust ports are set unilaterally at the fire source, the fire source is between the second export and the third one.

2.1 The Geometric Scale and Boundary Condition of the Tunnel Model

The study is based on Kong-Haiku-Mac undersea tunnel which is a two-way six-lane immersed tube tunnel. The research model is a two-way and eight-lane super wide tunnel with an ultra-wide section. A 1000-m part of tunnel is chosen to do numerical simulation calculation [24]:

- (1) Model scale: 1000 m (length) * 19.25 m (width) * 7.8 m (height). The cross-section area is 150.16 m².
- (2) Boundary condition: With the ignorance of influence of longitudinal wind, longitudinal air velocity is 0 m/s. Natural vent at the tunnel access is open with the relative pressure of 0 Pa. The initial environmental temperature in the tunnel is 17 °C. The total duration of the simulation is 1200 s.
- (3) Setup of exhaust ports: Exhaust ports are set at the location that is 1 m way from the downstream fire source, and the longitudinal interval between exhaust ports is 60 m. The exhaust port is sized at 4 m (length) * 1.5 m (width). The wind age area is 6 m². The air velocity at exhaust ports are listed at Table 1.

Heavy fire is simulated at the most adverse portion where smoke cannot spread to adjacent main tunnel. Hence, only the single-pole tunnel which is on fire is simulated. The three-dimension model of the tunnel is shown in Fig. 2

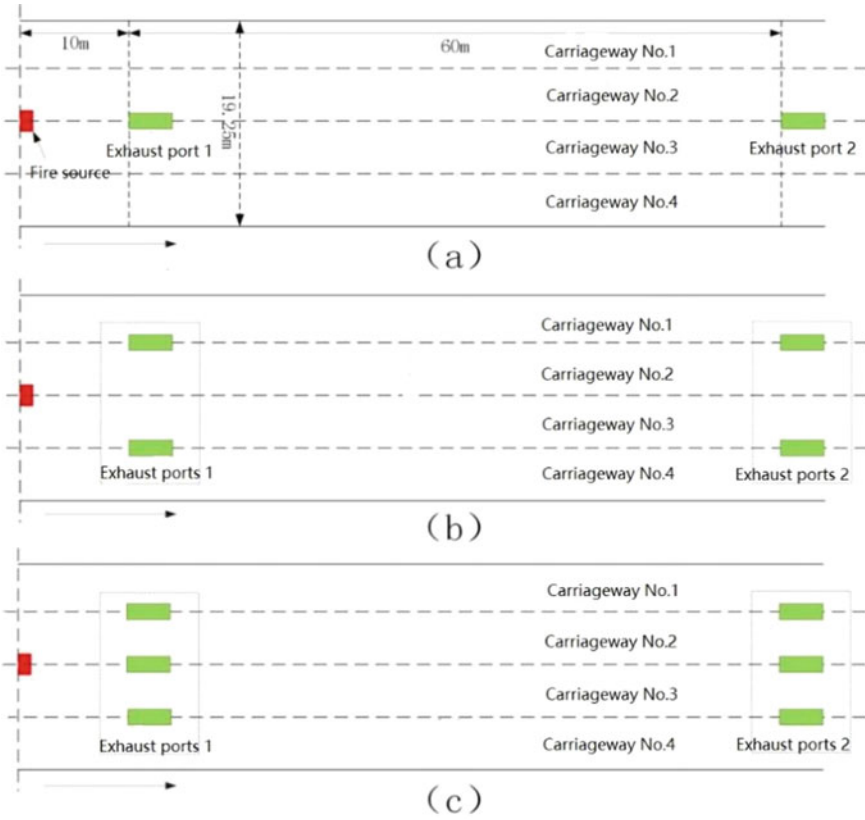


Fig. 1 Setup of exhaust port setting unilaterally at the fire source: **a** in one row; **b** in two row; **c** in three row

2.2 Fire Scenario

- (1) Fire source parameters: Fire source on the center-line of the tunnel is located at the boundary of the second and third traffic lanes. The fire source is 500 m away from the tunnel entrance near fire source upstream and 10 m away from the closest exhaust port to downstream fire source. The dimension of the fire source is 0.4 m high, 2 m wide and 2.5 m long with a volume of about 2 m³. The area of the upper surface of fire source is 5 m².
- (2) Combustion Model: The simulation of tunnel fire is based on diesel. The combustion model is to be C = 12, H = 23. The simulation takes 47 MJ/kg for the fuel heat value [25]. The value of C and H of the combustion model are set to be 12 and 23, respectively.
- (3) Fire scale: Large-scale fire model is simulated in this research, and fire source power is set to be 50 MW.

Table 1 Conditions of each case

Case No.	Setup modes of exhaust ports	Power of fire source	Air velocity at exhaust ports (m/s)	Theoretic volume of smoke exhaust (m ³ /s)	Positions of exhaust ports	Direction setting
1	Single row longitudinally	Large fire, 50 MW	9	216	Traffic line No. 2	On one side of the fire source
2	Double rows longitudinally		4.5	216	Traffic line No. 1, 3	
3	Three rows longitudinally		3	216	Traffic line No. 1, 2, 3	
4	Single row longitudinally		9	216	Traffic line No. 2	On both side of the fire source
5	Double rows longitudinally		4.5	216	Traffic line No. 1, 3	
6	Three rows longitudinally		3	216	Traffic line No. 1, 2, 3	

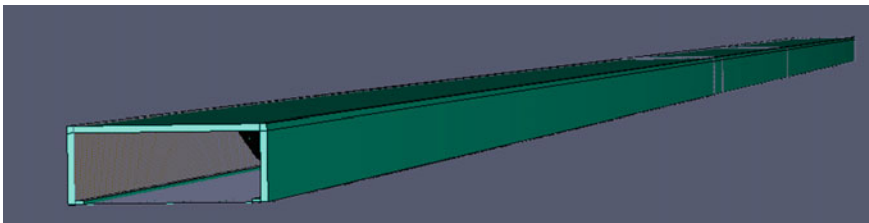


Fig. 2 The three-dimension model of the tunnel

- (4) Fire increasing modulus: It's shown in past researches that fire source power in the tunnel generally reaches the maximum within a very short time. Fire spreading at ultra-high speed is simulated in this study while ignoring fire attenuation stage. Heat Release Rate (HRR) changes according to the law of Eq. (1). The value of α is set to be 0.1876. When fire achieves stability, HRR remains basically constant [26]:

$$Q = \alpha t^2 \tag{1}$$

In this formula, Q represents HRR of which the unit is KW. α is the preferred coefficient of the equation, and t represents the fire-lasting time.

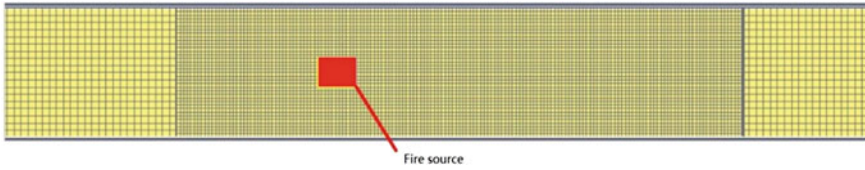


Fig. 3 The schematic of tunnel model grids

2.3 Grid Setup Methods

In numerical simulation, grid accuracy is an important factor affecting the computation result. The smaller the size of grid is used, the higher the precision of simulation will be. In consequence, a more accurate result will be obtained. However, a huge number of meshes will result in burdensome calculation and affect the efficiency of simulation seriously. Tunnel model for the simulation should be divided into computing grids with appropriate size. The research indicates that one-tenth of the fire characteristic diameter of fire source is used, which will be in more compliance with actual condition. The characteristic diameter of fire source D^* is defined as follows [27]:

$$D^* = \left[\frac{\dot{Q}}{\rho_{\infty} c_p T_{\infty} \sqrt{g}} \right]^{2/5} \quad (2)$$

with \dot{Q} representing the total heat release rate of fire source (KW), ρ_{∞} represents the density of hot air (kg/m^3). C_p represents specific heat at constant pressure (J/kg K). T_{∞} is the temperature of hot air (K), and g is the gravitational acceleration (m/s^2).

At $Q = 50$ MW, one-tenth of D^* is 0.25 m. The area near the fire source is the major region for generating smoke and that near exhaust ports for exhausting smoke. Others are smoke spread regions whose accuracy can be reduced appropriately. Thus, grid areas are divided into 2 zones: (1) smoke generation and exhaust zones (490–700 m), whose grid precision is $0.25 \text{ m} * 0.25 \text{ m} * 0.25 \text{ m}$; and (2) smoke spreading zone (0–490, 700–1000 m), whose grid precision is $0.5 \text{ m} * 0.5 \text{ m} * 0.5 \text{ m}$ (Fig. 3).

2.4 Setup Methods of Measurement Points

Layer partition: 100 monitoring devices (Layer Zoning Device) are longitudinally set at equal spacing along the tunnel centerline, the distance between which is 10 m.

Thermocouple: To monitor the temperature change at the safe height of the tunnel, 100 thermocouples are longitudinally set along the tunnel centerline with a distance of 2 m from the ground; and 100 measurement points are longitudinally set at equal

spacing along the tunnel centerline with a distance of 0.6 m from the vault, which is used to monitor the temperature change around the vault.

Flow measurement system: Heat-flow monitoring interface is set at the exhaust port, coordinate parameters of which are equal to those of the exhaust port.

CO₂: To monitor the mass flux of CO₂ around each exhaust port, Parameters such as MASS_FLUX Z are introduced to measure the mass flux in Z direction, and SPEC-ID is set as "carbon dioxide". Coordinate parameters are equal to those of the exhaust port.

3 Results and Discussion

3.1 Smoke Thickness

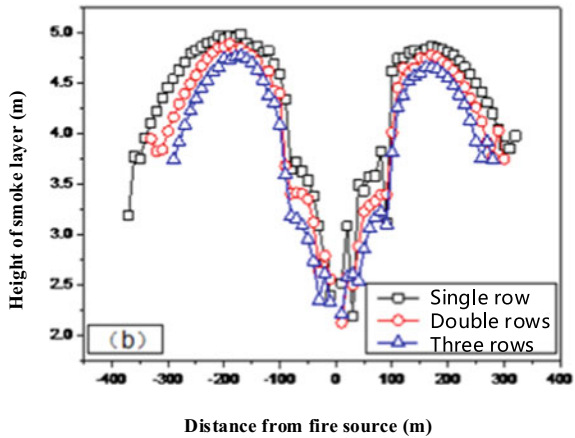
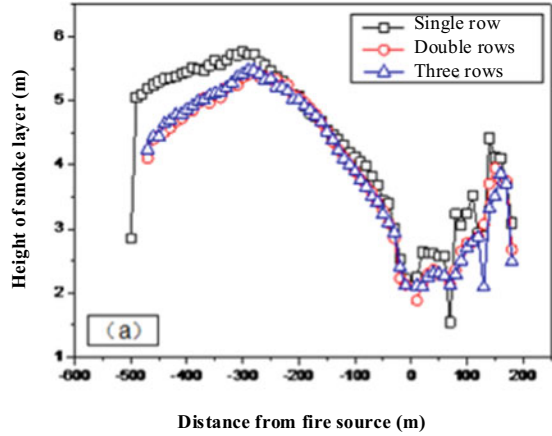
When the tunnel is on fire, people escaping will be harmed seriously due to smoke inhalation, which may cause harmful effects such as oxygen deficit, loss of muscle function, acute confusion and judgement recession. It's shown in concerning statistics of fire disaster that most victims die from inhalation of noxious and harmful gas such as CO.

Both upstream and downstream fire source, both two-column and three-column arrangement of exhaust ports have thicker smoke layer than initial single-column arrangement. When exhaust ports are set unilaterally, differences between smoke thickness of downstream fire source is more obvious on different conditions. In addition, unilateral-setting exhaust ports will generate smaller difference of smoke thickness between longitudinally three-column setting way and two-column. The smoke thickness is basically overlapped under the condition of setting exhaust ports longitudinally. When exhaust ports are set bilaterally, each additional column of exhaust ports cause evident decrease of smoke thickness. In case of longer smoke extraction area, a column of exhaust ports will correspond to moderately decreased thickness of smoke layer. However, if another column of exhaust ports is added, it won't achieve the effect of decreasing smoke thickness (Fig. 4).

3.2 Vault Maximum Temperature

When the tunnel is on fire, the temperature in tunnel in case of fire will increase rapidly. High-temperature smoke generated by fire and affected by buoyancy will cause excessive heat accumulation. Compared with other parts in the tunnel, the tunnel vault is more likely to be destroyed. The vault temperature is of greater importance to evaluate the efficiency of smoke ventilation. Figure 5 shows that the vault max temperature changes slightly when the quantity of exhaust ports is increased.

Fig. 4 Smoke thickness: **a** one-side arrangement of exhaust ports; **b** two-side arrangement of exhaust ports

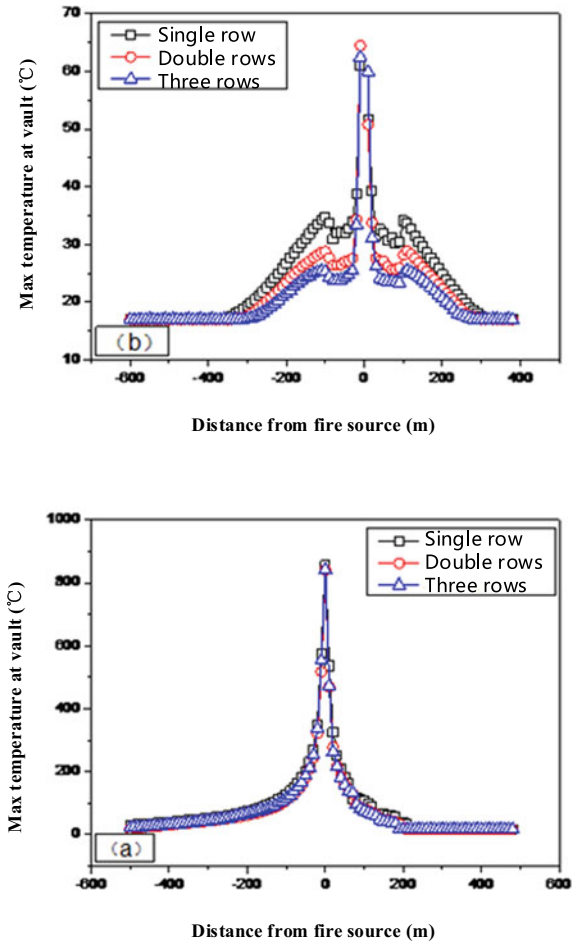


It also causes a less temperature difference in the smoke-exhausting area. The vault high-temperature curve is basically overlapped.

3.3 Temperature at the Safety Height

The harm that is induced by high-temperature smoke may be caused by direct contact and heat radiance. When exposed in an environment of 95 °C, people will have a symptom of dizziness only lasting for 1 min. People exposed in the environmental temperature of 120 °C for 1 min will get burnt. The temperature limitation for breathing is 131 °C. When environmental temperature exceeds 90 °C, people

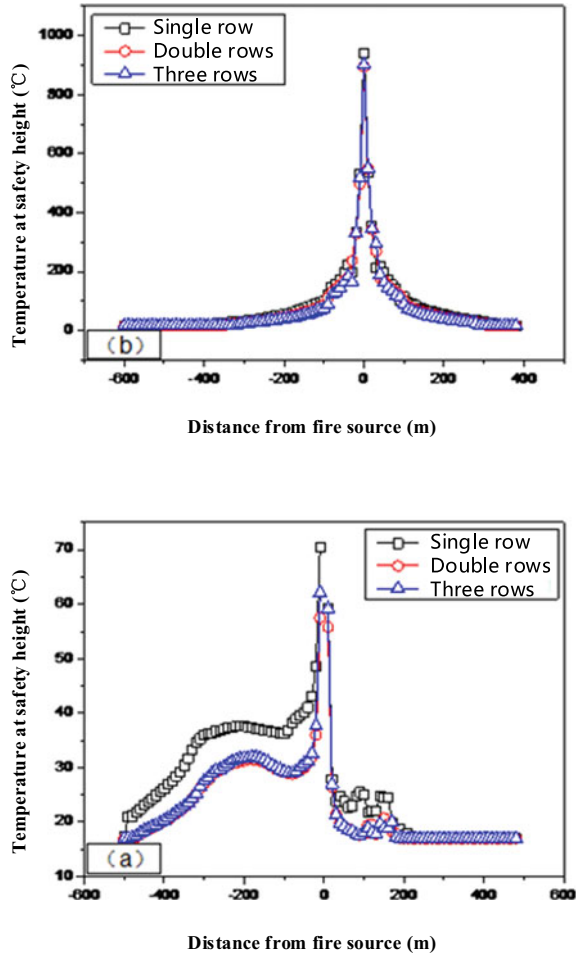
Fig. 5 Vault max temperature **a** one-side arrangement of exhaust ports; **b** two-side arrangement of exhaust ports



will be on incapacitation condition. Researches indicate that safety temperature for evacuation is usually lower than 66 °C.

The temperature at the safety height above fire source is not compared with that in other place, for it is rather higher. As shown in Fig. 6, the smoke-exhausting region among both upstream and downstream fire source, the temperature in case 1 and case 2 is much lower than that in case 3. There is no obvious difference between case 2 and case 3. The longitudinal temperature curves of both are basically overlapped.

Fig. 6 Temperature at the safety height **a** one-side arrangement of exhaust ports; **b** two-side arrangement of exhaust ports



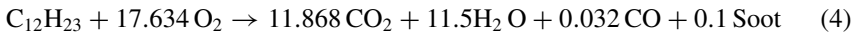
3.4 Smoke Extraction Efficiency

In semi-transverse centralized smoke extraction system, smoke extraction efficiency, a quantitative index of smoke-exhausting effect, can make a quantitative analysis on smoke-exhausting effect. The smoke-exhausting efficiency falls into two categories: smoke-exhausting efficiency of exhaust ports and smoke-exhausting efficiency of smoke extraction system, which is given by Eq. (3) [28]:

$$\eta = \frac{m_e}{m_p} \times 100\% = \frac{\sum m_e}{m_p} \times 100\% \tag{3}$$

where m_e denotes the total flow gas volume from exhaust ports and m_p represents that generated by fire source.

Smoke generated during combustion is always with air entrainment. It's very difficult to determine the smoke thickness in case of only little smoke generated at the exhaust port. As a result, it's also difficult to determine generation quantity of smoke and flue gas volume by numerical simulation. Due to the selected large tunnel model, fuel can be regarded to be in complete combustion with a product of CO_2 . The amount of CO_2 produced by combustion represents for flow gas amount and that extracted through all exhaust ports represents extracted smoke volume. The equation for complete combustion of diesel is as follows:



The CO_2 production per unit time is 3.32 kg/s, which is calculated from the reaction equation and taking heat of combustion for 47 MJ/kg [9].

As in Fig. 6a, with a one-side setting way of exhaust ports, attenuation degree of smoke-exhausting efficiency of first three continuous groups of smoke vents is lower. An evident attenuation of smoke extraction efficiency only occurs until the forth group of smoke vents. Smoke extraction efficiencies of four groups of exhaust ports decreases progressively in either case 2 or case 3. The 1st group of exhaust ports in both case 2 and case 3 have higher smoke-exhausting efficiency than that is in case 1.

After calculation, the overall smoke-exhausting efficiency of four groups of exhaust ports in case 1 is 40.63%, in case 2 is 37.06%, and that in case 3 is 37.86%. The overall smoke-exhausting efficiency in case 1 is higher than that in other two cases. The smoke-exhausting efficiency of the 1st group of exhaust ports suggests that the efficiency will increase with the increasing number of exhaust ports in the transverse direction. It's feasible to compare the efficiency of four groups of exhaust ports after changing the location of fire source.

As is shown in Fig. 6b, with single-row exhaust ports, efficiency of exhaust ports (port2 and port3) that are closer to the fire source are higher than those of exhaust ports (port1 and port4) farther away from fire source. With two-column exhaust ports, efficiency of first three exhaust ports are closer to each other, while the efficiency of port4 declines evidently. Moreover, with three-column exhaust ports, exhaust ports (port2 and port3) that are on both sides near fire source have much larger efficiency than those (port1 and port4) farther away from fire source.

It's calculated that the overall smoke-exhausting efficiency of four groups of exhaust ports in single row is 56.42%, that of four groups of exhaust ports in two rows is 60.43%, and that of four groups of three-row exhaust ports is 60.43%. The efficiency of both two-row and three-row exhaust ports are higher than that of single-row ones. As a result, the reason why the overall smoke-exhausting efficiency of exhaust ports in single row is always higher than that of ports in two and three rows is that a great deal of smoke is exhausted from the first group of exhaust ports near the fire source.

4 Conclusions

In this research, the smoke control performance of top centralized smoke exhaust strategy is investigated in the ultra-wide immersed tube tunnel.

- (1) For ultra-wide tunnel, it can efficiently reduce the smoke thickness as well as the vault maximum temperature and the temperature at safety height by increasing the number of exhaust ports along the transverse direction.
- (2) The smoke exhaust efficiency is significantly improved by increasing the number of exhaust ports along the transverse direction. Thus, it's strongly recommended to adopt multi-row exhaust ports rather than one-single-row exhaust ports.
- (3) The smoke exhaust efficiency of the exhaust port is obviously reduced with its distance far away from the fire. Consequently, if permitted, it's strongly recommended to reduce the spacing between exhaust ports in smoke control system design.

Author Contributions Y. L. (Ying Li) and K. T. (Kaixuan Tang) conceived the original ideal; Y. L. (Ying Li) wrote and edited the manuscript; and S. M. (Shenglin Mu), X. N. (Xiaoyang Ni), and Z. C. (Zhongchen Niu) supervised the study. All authors have read and agreed to the published version of the manuscript.

Funding This research was funded by the National Natural Science Foundation of China (Nos. 51607078).

Conflicts of Interest The authors declare no conflict of interest.

References

1. Zhu H, Shen Y, Yan Z et al (2016) A numerical study on the feasibility and efficiency of point smoke extraction strategies in large cross-section shield tunnel fires using CFD modeling. *J Loss Prev Process Ind* 44:158–170
2. Chow WK, Gao Y, Zhao JH et al (2016) A study on tilted tunnel fire under natural ventilation. *Fire Saf J* 81:44–57
3. Oka Y, Oka H (2016) Velocity and temperature attenuation of a ceiling-jet along a horizontal tunnel with a flat ceiling and natural ventilation. *Tunnell Undergr Space Technol Incorpor Trenchless Technol Res* 56:79–89
4. Hu LH, Fong NK, Yang LZ et al (2007) Modeling fire-induced smoke spread and carbon monoxide transportation in a long channel: fire dynamics simulator comparisons with measured data. *J Hazard Mater* 140:293–298
5. Furitsu Y, Koichi O, Takayoshi O (2004) Fire protection for TBM shield tunnel lining. In: *Tunnelling and underground space technology incorporating trenchless technology research*, p 14
6. Wang HY (2009) Prediction of soot and carbon monoxide production in a ventilated tunnel fire by using a computer simulation. *Fire Saf J* 44:394–406

7. Hu LH, Fong NK, Yang LZ et al (140) Modeling fire-induced smoke spread and carbon monoxide transportation in a long channel: fire dynamics simulator comparisons with measured data. *J Hazardous Mater* 140(1–2):293–298
8. Nordmark A (1998) Fire and life safety for underground facilities: present status of fire and life safety principles related to underground facilities. *Tunn Undergr Space Technol* 13:217–269
9. Wang Q, Tang F, Li LJ et al (2016) Large eddy simulation on the effect of smoke exhaust openings arrangement on the smoke spread in tunnel fires. *Proc Eng* 135:308–314
10. Liu Y, Cassady S (2014) A modified critical velocity for road tunnel fire smoke management with dedicated smoke extraction configuration. *Case Stud Fire Safety* 2:16–27
11. Zhisheng X, Yangyang L, Rui K et al (2014) Study on the reasonable smoke exhaust rate of the cross range exhaust duct in double-layer shield tunnel. *Proc Eng* 84:506–513
12. Wang Z, Chen B (2011) Discussion on Jiaozhou bay submarine tunnel smoke control design. *Fire Sci Technol*
13. Chen JJ, Fang Z, Yuan J (2013) Numerical simulation about smoke vent arrangement influence on smoke control in double-decked tunnel. *Proc Eng*:48–55
14. Ingason H, Li YZ (2010) Model scale tunnel fire tests with longitudinal ventilation. *Fire Saf J* 8:371–384
15. Lv Y, Cao B, Zhang LJ et al (2013) Numerical simulation on the influence of double-jet air curtain of smoke-preventing in tunnel fires. *Adv Mater Res* 805:534–1537
16. Buchanan AH, Science IAFF (1999) Implementation of performance-based fire codes. *Fire Saf J* 32:377–383
17. Chen JJ, Fang Z, Yuan J (2013) Numerical simulation about smoke vent arrangement influence on smoke control in double-decked tunnel. *Proc Eng* 52:48–55
18. Zhong W, Ji J, Yang JP et al (2010) Experimental and simulation study on the influence of smoke vent arrangement on smoke control effect in subway station. *Eng Mech*
19. Zhu J, Wu JB, Zhang ZG et al (2014) A study on the technique for fire smoke control in urban traffic link tunnel. *Appl Mech Mater* 640:1117–1127
20. Lin P, Lo SM, Li T (2014) Numerical study on the impact of gradient on semi-transverse smoke control system in tunnel. *Tunnell Undergr Space Technol Incorpor Trenchless Technol Res* 44:68–79
21. US Department of Commerce N. *Fire Dynamics Simulator (Version 2)*, Technical Reference Guide. National Institute of Standards & Technology
22. Li JSM, Chow WK (2003) Numerical studies on performance evaluation of tunnel ventilation safety systems. *Tunn Undergr Space Technol* 18:435–452
23. Vega MG et al (2008) Numerical 3D simulation of a longitudinal ventilation system: memorial tunnel case. *Tunnell Undergr Space Technol Incorporating Trenchless Technol Res* 23:539–551
24. Liu S, Jiang Y, Chen J et al (2016) Full-scale experimental research on the effect of smoke exhaust strategies on efficiency of point extraction system in an underwater tunnel. *Proc Eng* 166:362–372
25. Şahan N, Fois M, Paksoy H (2015) Improving thermal conductivity phase change materials—a study of paraffin nanomagnetite composites. *Sol Energy Mater Sol Cells* 137:61–67
26. Drysdale D (2011) *An introduction to fire dynamics*, 3rd edn.
27. Kim E, Woycheese JP, Dembsey NA (2008) Fire dynamics simulator (version 4.0) simulation for tunnel fire scenarios with forced, transient, longitudinal ventilation flows. *Fire Technol* 44:137–166
28. Cong HY, Wang XS, Zhu P et al (2017) Improvement in smoke extraction efficiency by natural ventilation through a board-coupled shaft during tunnel fires. *Appl Therm Eng* 115:127–137

Research on Safety Management of Inland River Dangerous Chemicals Transportation System Based on System Dynamics and Principal Component Analysis



Yu Zhao and Zhiyong Lv

Abstract This paper proposes a new method to study the safety of inland river hazardous chemical transportation under dynamic conditions based on the principles of system dynamics by analyzing the related theories of hazardous chemical transportation and comparing the traditional static hazard assessment methods. The system dynamics model of the inland river hazardous chemicals transportation system is constructed in this paper. According to Vensim_PLE, relevant data are obtained by changing the growth rate of the safety input of various factors. Using the principal component analysis (PSC) analysis data to verify the accuracy of the model and study the impact of various factors on the safety of the transportation system. The result shows that human factors and management factors have a greater influence on the safety of the transportation system. It is important to change the input indicators of human factors and management factors to improve safety.

Keywords Transportation of inland river dangerous chemicals · Safety management · System dynamics

1 Introduction

Dangerous chemicals are extremely uncertain and dangerous during production and transportation [1]. There is a huge amount of energy stored in hazardous chemicals. People use energy to achieve production purposes. If the energy is out of control it will inevitably cause accidents.

Research on the transportation of hazardous chemicals is mostly based on land transportation [2]. Compared with road and rail transportation, inland river transportation has the advantages of large transportation volume environmental protection and low transportation cost.

Y. Zhao · Z. Lv (✉)

National Water Transport Safety Engineering Technology Research Center, Wuhan University of Technology, Wuhan 430000, China

e-mail: benlzy@whut.edu.cn

In recent years, with the development of direct transportation between the river and the sea, the number of dangerous chemical vessels has increased rapidly. However, there are few studies on water transportation.

System dynamics has inherent advantages for solving complex problems. At present, system dynamics is widely used in the research of urban economic development enterprise management energy issues and other fields. And system dynamics solves problems by establishing mathematical models.

2 System Dynamics Model of Inland River Hazardous Chemicals Transportation System

According to the reference [3], the influencing factors of hazardous chemicals road transportation are people, mechanical equipment, environment, management, and hazardous chemicals. This work also considers the factors that affect the safety of the ship. So the factors affecting the safety of the inland river chemical transportation system are roughly divided into six parts: ship factor, chemical factor, human factor, equipment factor, environmental factor, and management factor.

In order to construct the dynamic flow chart and system dynamics equation of inland river hazardous chemicals transportation system, the following representative relevant variables were selected:

System: in this article refers to the system security level indicator (FTgt).

Ship factors: ship factor safety index (Tgt1), ship factor contribution rate to system safety level (C1), ship factor influence rate (referring to positive impact) (I1), ship factor attenuation rate (D1), ship factor safety level added value (E1), the growth rate of ship factor safety investment (Z1), ship factor safety input level (P1).

Chemical factors: Chemical factor safety index (Tgt2), chemical factor contribution rate to system safety level (C2), chemical factor influence rate (referring to positive impact) (I2), chemical factor attenuation rate (D2), the added value of chemical factor safety level (E2), a growth rate of chemical factor safety input (Z2), chemical factor safety input level (P2).

Human factors: Human factor safety index (Tgt3), human factor contribution rate to system security level (C3), human factor influence rate (referring to positive impact) (I3), human factor attenuation rate (D3), human factor safety level Added value (E3), human factor safety investment growth rate (Z3), human factor safety investment level (P3).

Equipment factors: equipment factor safety index (Tgt4), equipment factor contribution rate to system safety level (C4), equipment factor influence rate (referring to positive influence) (I4), equipment factor attenuation rate (D4), equipment factor safety level Added value (E4), equipment factor safety investment growth rate (Z4), equipment factor safety investment level (P4).

Environmental factors: environmental factor safety index (Tgt5), environmental factor contribution rate to system safety level (C5), environmental factor impact rate

(referring to positive impact) (I5), environmental factor attenuation rate (D5), environmental factor safety level Increased value (E5), the growth rate of environmental factor safety investment (Z5), level of environmental factor safety investment (P5).

Management factors: management factor safety index (Tgt6), management factor contribution rate to system security level (C6), management factor influence rate (referring to positive influence) (I6), management factor decay rate (D6), management factor safety level Added value (E6), the growth rate of safety investment of management factors (Z6), level of safety investment of management factors (P6).

3 System Dynamics Simulation

In this work, we construct a causality diagram of the main information of the system as shown in Fig. 1 according to the six main influencing factors analyzed in the second part. Positive polarity means positive feedback and negative means negative feedback [4].

The system dynamics equations of the inland river hazardous chemicals transportation system are as follows:

$$\begin{aligned}
 Tgt1.X &= Tgt1.G + (M) \times Tgt1.G \times (I1.GX - D1.GX); \\
 Tgt2.X &= Tgt2.G + (M) \times Tgt2.G \times (I2.GX - D2.GX); \\
 Tgt3.X &= Tgt3.G + (M) \times Tgt3.G \times (I3.GX - D3.GX); \\
 Tgt4.X &= Tgt4.G + (M) \times Tgt4.G \times (I4.GX - D4.GX); \\
 Tgt5.X &= Tgt5.G + (M) \times Tgt5.G \times (I5.GX - D5.GX); \\
 Tgt6.X &= Tgt6.G + (M) \times Tgt6.G \times (I6.GX - D6.GX); \\
 P1.X &= P1.G + (M) \times P1.G \times (Z1.GX);
 \end{aligned}$$

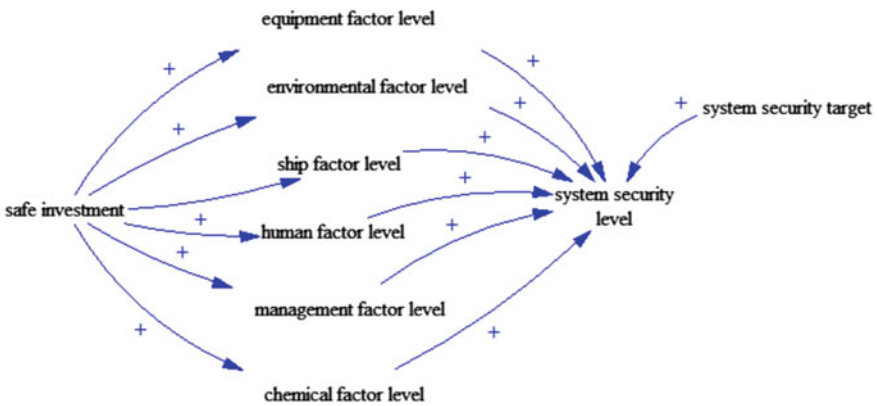


Fig. 1 System causality diagram

$$\begin{aligned}
 P2.X &= P2.G + (M) \times P2.G \times (Z2.GX); \\
 P3.X &= P3.G + (M) \times P3.G \times (Z3.GX); \\
 P4.X &= P4.G + (M) \times P4.G \times (Z4.GX); \\
 P5.X &= P5.G + (M) \times P5.G \times (Z5.GX); \\
 P6.X &= P6.G + (M) \times P6.G \times (Z6.GX); \\
 STgt.X &= C1 \times Tgt1.X + C2 \times Tgt2.X + C3 \\
 &\quad \times Tgt3.X + C4 \times Tgt4.X + C5 \\
 &\quad \times Tgt5.X + C6 \times Tgt6.X \\
 &\quad (C1 + C2 + C3 + C4 + C5 + C6 = 1); \\
 I1.XJ &= Z1 \times (P1.X); I2.XJ = Z2 \times (P2.X); \\
 I3.XJ &= Z3 \times (P3.X); I4.XJ = Z4 \times (P4.X); \\
 I5.XJ &= Z5 \times (P5.X); I6.XJ = Z6 \times (P6.X);
 \end{aligned}$$

The system dynamics simulation diagram is shown in Fig. 2.

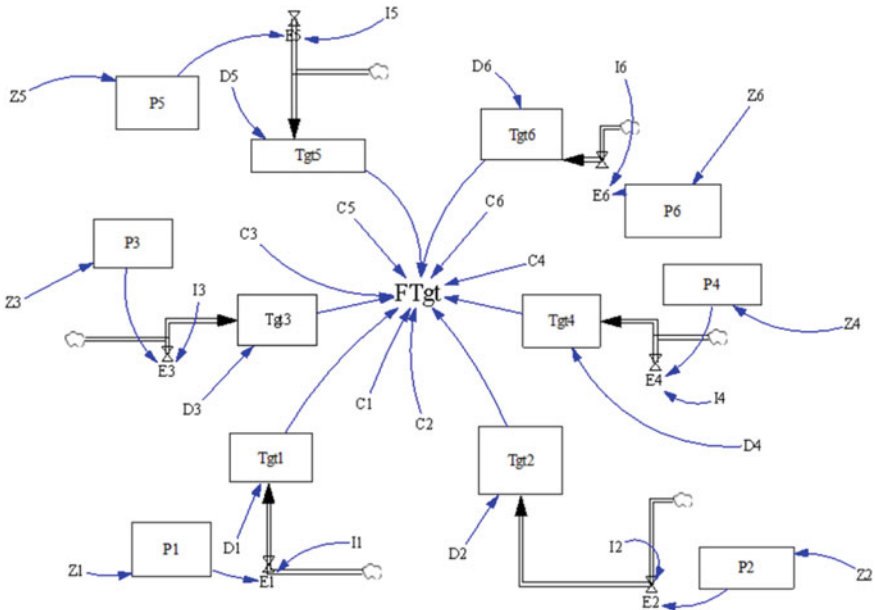


Fig. 2 System dynamics simulation flow

4 Simulation Input and Result Analysis

This article refers to references [4–6] and other documents and makes minor adjustments. Finally, the initial values are set as follows, and the numerical order corresponds to the human factor, equipment factor, and environmental factor, management factor, shipping factor, chemical factor. The initial value of the safety level index of each factor is taken as (75, 80, 80, 75, 80, 75). The initial value of the safety input to each factor is 1 and the contribution rate of each factor to the system safety level is (0.4, 0.05, 0.1, 0.3, 0.1, 0.05). The simulation step duration is 1 month. The decay rate of the safety level of each factor is (0.04, 0.01, 0.01, 0.04, 0.005, 0.005) and the growth rate of the safety input of each factor is taken as (0.2, 0.2, 0.2, 0.2, 0.2, 0.2). The influence coefficient of each factor (referring to the positive influence on the safety level of this factor) is (0.012, 0.015, 0.006, 0.02, 0.005, 0.006) and the target index value of the system safety level is set as 99.

Firstly, input the dynamic equations and the initial values in Vensim_PLE. Secondly, run the simulation. Finally, get the value change of the system safety level as shown in Table 1.

It can be seen from these data that the increase of safety investment will cause an increase in various factors and will surely cause the rise of the system safety level. It is necessary to further increase the safety investment of each factor to further increase the safety level of the system. However, the safety level value will show a downward trend within a certain period and then gradually increase when increasing the safety investment to increase the system safety level. So the system can be dynamically evaluated. From Table 1, the value of the safety level index is 75.75 in the first month. It means that the safety state of the system is in a bad state and the safety level index of the system shows a downward trend. Because the initial input value is low and the decay rate of each factor is high. It also shows a downward trend after further increasing the safety input. The reason is that the system has a detention effect. With the further increase of the safety input of each factor, the safety level index of the system shows an upward trend due to the increase of the base number. The safety level index has exceeded the design index at the beginning of December so it can be considered that the safety status of the system is in a good state. At this time, the safety investment can be appropriately reduced according to the actual situation. Table 2 shows the values of safety level indicators of each factor.

The input growth rate of each factor will be changed in turn to further analyze the degree of influence of the input level of each factor on the system security level. For the first time, the input growth rate of chemical factors is changed to 0.4 and the rest unchanged. For the second time, environmental factors are changed and the rest unchanged. From the third time to the sixth time, human factors, equipment factors, management factors, and ship factors are successively changed. The values of the changed system safety level indicators are shown in Table 3 and Fig. 3.

The input growth rate of each factor will be changed in turn to further analyze the degree of influence of the input level of each factor on the system security level. For the first time, the input growth rate of chemical factors is changed to 0.4 and the

Table 1 System security level indicators

Time/month	1	2	3	4	5	6	7	8	9	10	11	12
Numerical value	75.75	73.41	71.17	69.05	67.08	65.35	64.03	63.52	64.69	69.46	82.01	111.5

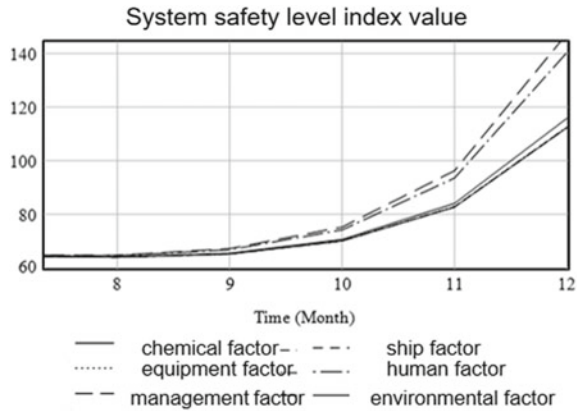
Table 2 Values of safety level indicator

Time/month	1	2	3	4	5	6	7	8	9	10	11	12
Human factors	75	72.01	69.15	66.44	63.91	61.64	59.79	58.76	63.61	75.55	92.00	104.40
Environmental factors	75	74.25	73.52	72.82	72.15	71.57	71.16	71.13	71.92	74.49	80.99	96.12
Chemical factors	75	74.62	74.26	73.91	73.57	73.28	73.07	73.01	73.26	74.12	76.21	80.74
Ship factors	80	79.60	79.21	78.84	78.49	78.21	78.10	77.99	78.37	79.51	82.18	87.91
Management factors	75	72.02	69.18	66.51	64.06	61.97	60.53	60.36	62.94	71.40	92.69	142.11
Device factors	80	79.21	78.45	77.74	77.11	76.61	76.42	76.80	78.83	74.49	82.16	108.07

Table 3 Values of the system safety level

Time/month	1	2	3	4	5	6	7	8	9	10	11	12
Chemical factors	75.75	73.49	71.32	69.28	67.39	65.72	64.47	64.02	65.26	70.10	82.73	112.32
Environmental factors	75.75	73.49	71.32	69.28	67.38	65.73	64.49	64.09	65.45	70.60	84.07	115.82
Human factors	75.75	73.49	71.33	69.29	67.40	65.79	64.66	64.58	66.80	74.21	93.54	140.34
Device factors	75.75	73.49	71.33	69.28	67.39	65.72	64.47	64.04	65.29	70.16	82.84	112.54
Management factors	75.75	73.49	71.33	69.29	67.41	65.80	64.71	64.72	67.19	75.25	96.25	147.37
Ship factors	75.75	73.49	71.33	69.28	67.39	65.72	64.47	64.03	65.28	70.13	82.79	112.44

Fig. 3 System security level indicator composite map



rest unchanged. For the second time, environmental factors are changed and the rest unchanged. From the third time to the sixth time, human factors, equipment factors, management factors, and ship factors are successively changed. The values of the changed system safety level indicators are shown in Table 3 and Fig. 3 (Tables 4 and 5).

Calculate the percentage of the total variance explained by each principal component.

It can be seen from the Table 6 that the proportion of the first principal component is 99.6135% so the first principal component can fully reflect the original data information. The percentage of each element in the first principal component is shown in Table 7 and described by the Pareto chart.

Table 4 Principal components

0.3193	- 0.4033	- 0.3140	0.7235	0.3360	- 0.0200
0.3209	- 0.3947	0.6681	- 0.0943	0.0169	- 0.5345
0.5596	0.5478	0.1204	0.3446	- 0.5035	- 0.5345
0.3202	- 0.3984	0.2274	- 0.1789	- 0.1377	0.7976
0.5114	0.3568	-0.0821	- 0.3771	0.6792	0.0289
0.3433	- 0.3088	-0.6182	- 0.4179	- 0.3912	- 0.2772

Table 5 Variances

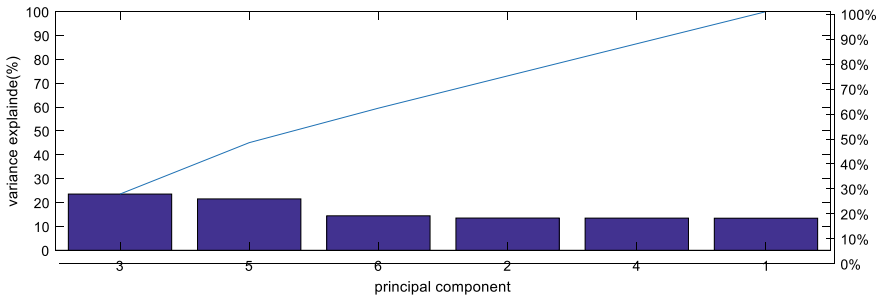
1.7437	0.0068	0.0000	0.0000	0.0000	0.0000
--------	--------	--------	--------	--------	--------

Table 6 Percentage of total variance explained by each principal component

99.6135	0.3865	0.0000	0.0000	0.0000	0.0000
---------	--------	--------	--------	--------	--------

Table 7 Percentage of each element in the first principal component

13.4459	13.5133	23.5651	13.4838	21.5354	14.4566
---------	---------	---------	---------	---------	---------

**Fig. 4** Pareto chart

It can be seen from Fig. 4 that the third element has the highest proportion, followed by the fifth element, and the other elements have a relatively small percentage. According to the data in Table 3, the third element corresponds to human factors and the fifth element corresponds to management factors. Therefore, the first principal component contributes to the system security level. It is necessary to strengthen the inspection of management defects increase training efforts and control human error to improve the safety management of hazardous chemical transportation systems.

5 Conclusion

Based on the hazard evaluation method and the theory of system dynamics method, this work analyzes the influence of the input growth rate of the ship factor chemical factor human factor management factor equipment factor and environment factor on the system safety level by using the principal component analysis method. It provides a new theory and practice method of system dynamics safety assessment for inland river hazardous chemical transportation. Through the analysis of this paper, we get the following conclusions:

- (1) SD modeling method can connect the relationship and interaction between different parameters in the system through its related feedback loop. The analysis process is simple and direct. It improves the analysis efficiency.
- (2) It is easy to find out which factor is most important by improving the traditional static source static assessment method through principal component analysis. Through analysis, we can find that the two elements of personnel and management are very important in improving the safety management process.
- (3) The established system dynamics model uses the contribution rate to connect the value of each input factor safety level index with the value of the total

system safety level. It adjusts the safety management level by adjusting the input growth of each factor.

In summary, through the safety analysis of the inland river hazardous chemicals shipping system, human factors and management factors could change the safety trend when environmental safety deteriorates. The use of system dynamics methods could improve the safety effect and provide a scientific basis for safety practice.

References

1. Ning S, Jingjing W, Ying Z (2019) Safety risk analysis and control of road transportation of hazardous chemicals. *Logist Technol* 42(07):97–98
2. Jia H, Li C, Wang Z et al (2018) Risk evaluation of hazardous chemicals transportation based on fuzzy comprehensive evaluation. *Chem Eng Trans* 71:271–276
3. Songbin D, Rong S, Heping S (2006) Research on safety management of air traffic system based on system dynamics. *J Transp Eng Inf* 2006(04):1–6+99
4. Fan L, Lian S (2004) Early warning management of air traffic disasters. Hebei Science and Technology Press, Shijiazhuang, pp 342–349
5. Tian J, Zhu M (2008) System dynamics in calculation on probability of spares sufficiency. *J Syst Simul (S1004-731X)* 20(1):299–302
6. Prochaska JD, Kim H, Buschmann RN et al (2019) The utility of a system dynamics approach for understanding cumulative health risk from exposure to environmental hazards. *Environ Res* 172:462–469

Mechanical Properties of Concrete Under the Alternate Carbonation and Freezing–Thawing Actions



Jin Ran, Qi Feng, Jinxi Zhang, Lin Wang, and Honglin Cao

Abstract Three concrete mixes with different water-to-binder ratios and dosages of cement replacement materials (fly ash and slag powder) were exposed to alternate carbonation and freezing–thawing actions. The compressive strength, dynamic elastic modulus and mass of all specimens were measured and analyzed, also compared to the results obtained from the carbonation-only and freezing–thawing-only cases. Test data show that the relative compressive strength, the relative dynamic elastic modulus and the mass of concrete specimens exposed to carbonation-only cases increase with the carbonation time due to sufficient carbonation reaction, whereas they decrease with the corresponding testing time whenever they are exposed to freezing–thawing-only cases or alternate regimes. Additionally, under the same freezing–thawing numbers, the concrete specimens exposed to alternate regimes have slightly higher relative compressive strength, lower relative dynamic elastic modulus and larger mass loss than that exposed to freezing–thawing-only cases. Based on the obtained dynamic elastic modulus from the above different environmental cases, a damage mechanics fitting model of concrete considering alternate carbonation and freezing–thawing actions were established with high reliability.

Keywords Concrete · Alternate carbonation and freezing–thawing actions · Relative compressive strength · Relative dynamic elastic modulus · Mass · Damage mechanics fitting model

1 Introduction

In concrete industry, more and more researchers realize that the deterioration of concrete cannot be really reflected just by simulating the single environmental factor

J. Ran (✉) · Q. Feng · L. Wang · H. Cao
Shandong Transportation Institute, Jinan 250102, P. R. China
e-mail: ranjin@sdjtky.cn

J. Zhang
Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, P. R. China

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_8

in laboratory [1–3]. Therefore, some research efforts have begun investigating the concrete deterioration process considering the combined actions of various factors. Ramezani-pour et al. developed an environmental chamber to simulate tide cycles, and investigated the microscopic and mechanical properties of ordinary concrete exposed to CO₂ gas, saline water, and the combination of CO₂ gas and saline water by an accelerated test method [4]. David et al. proposed a numerical approach TransChlor to stimulate the realistic microclimatic conditions, and investigated the transportation phenomena of various substances in concrete [5, 6].

In cold regions, the reinforced concrete structures always subject to the coupling damage actions of freezing–thawing and CO₂ gas in their surrounding environment. The material and structural integrity of concrete are seriously affected by the actions of freezing–thawing [7], and CO₂ gas destroys the high pH environment of concrete pore solution which protects the steel bars of reinforced concrete from corrosion [8]. Therefore, the durability of concrete subjected to the combined actions of carbonation and freezing–thawing is important and needs to be studied.

Xiao et al. measured the mass loss rate and relative dynamic elastic modulus of concrete specimens subjected to alternate freezing–thawing and carbonation actions, showing that their combined actions accelerated the concrete damage [9]. Zhao et al. investigated the durability and pore structure of concrete exposed to freezing–thawing and carbonation in turn, showing that carbonation improved the freezing resistance of concrete in the early period, but the trend was reversed in the later period [10]. Jiang et al. investigated the frost and carbonation resistance of concrete with different water to cement ratios and fly ash contents under alternate freezing–thawing and carbonation actions, showing that the carbonation depth of concrete under the alternate actions didn't follow a square root-of-time trend, also increased faster with alternate times [11].

From the above research efforts, it was found that the present experimental approaches on the alternate actions between freezing–thawing and carbonation were significant different, such as the sequences of test procedures, the duration time of each test, and the conditions of each test etc. Generally, the freezing–thawing test was always conducted at first in most previous research works, i.e. the concrete specimens had already been damaged before the carbonation test. This paper proposed an opposite experimental approach, i.e. the carbonation test was conducted at first, and then followed by the freezing–thawing test to simulate the coupling effect of carbonation and freezing–thawing. It is believed that the new sequence of testing procedure reflects the real environment concrete experiences, since the hydration procedure and calcium hydroxide content at the surface of concrete trigger carbonation before freezing–thawing happens [12]. Moreover, the damage mechanics fitting model of concrete exposed to alternate carbonation and freezing–thawing actions is attempted to propose based on the relationship developed from the test results of single environmental factors.

Table 1 Mix proportions of concrete

Code	m_w/m_b	Sand ratio (%)	Mix proportions (kg/m ³)					
			Cement	FA	SP	Water	Sand	Stone
C1	0.44	42.7	340	0	0	150	850	1140
C2	0.42	42.9	293	52	0	145	840	1120
C3	0.40	42.8	253	55	55	145	832	1110

FA fly ash, SP slag powder

2 Materials and Experimental Methods

2.1 Materials

Commercially available ordinary Portland cement (PO 42.5R) is used as binder in concrete, which satisfies the Chinese standard GB 175–2007 specification. Fly ash (FA) conforms to the category of Class F Grade I in the Chinese standard GB/T 1596-2005 specification. Slag powder conforms to the category of S95 in the Chinese standard GB/T 18046-2008 specification. The fine aggregate is unwashed ordinary natural sand, and its fineness modulus is 2.6, equivalent to medium sand, and conforms to gradation zone 2 in the Chinese standard GB/T 14684-2011 specification. The coarse aggregate is unwashed crushed limestone, which nominal size is 5–25 mm and apparent density is 2.821 g/cm³, conforms to the requirements of coarse aggregate in the Chinese standard GB 14685-2011 specification. The chemical admixture is polycarboxylates high performance water reducing agent. The mixing water is ordinary tap water.

2.2 Mix Proportions

Three concrete mixes were used to investigate the effect of alternate carbonation and freezing–thawing actions on mechanical properties of concrete, their mix proportions were shown in Table 1.

2.3 Casting and Curing

Different size concrete specimens were casted in the moulds according to the above mix proportions. The cube specimens measuring 100 × 100 × 100 mm were used to determine the compressive strength, and the prism specimens measuring 100 × 100 × 400 mm were used to measure the mass and dynamic elastic modulus. These specimens were demolded after 24 h, and all of them were left to cure in standard

curing chamber (20 ± 2 °C and more than 95% relative humidity) for 56 days to start the relevant tests.

2.4 Experimental Methods

Based on the test procedures on carbonation-only and freezing–thawing-only cases in the Chinese standard GB/T 50082-2009 specification, the experimental approaches on alternate carbonation and freezing–thawing actions were as follows: the carbonation test was conducted first after the concrete specimens were dried in an oven at 60 °C for 48 h, and exposed to the environment of $20 \pm 3\%$ CO₂ concentration, 20 ± 2 °C temperature and $70 \pm 5\%$ relative humidity for 7 days. Then, the specimens after carbonation test were immersed in water at 20 ± 2 °C for 4 days, then were left in the apparatus of freezing–thawing cycles, where the lowest freezing temperature was -18 ± 2 °C, and the highest thawing temperature was 5 ± 2 °C. The time of every freezing–thawing cycle was about 3.5 h, and the corresponding cycle numbers were 50 (equivalent to 7 days). Then the compressive strength, mass and dynamic elastic modulus of concrete specimens were measured at each of the 6 times alternate.

3 Results and Discussions

3.1 Relative Compressive Strength of Concrete Exposed to Different Environmental Actions

Figure 1 showed that the relative compressive strength of concrete specimens exposed to carbonation-only cases increased with the carbonation time, largely due to carbonation led to the formation of important carbonation product (CaCO₃) [4]. However, the relative compressive strength of concrete specimens exposed to freezing–thawing-only cases or alternate carbonation and freezing–thawing actions decreased with the corresponding testing time. Under the same number of freezing–thawing cycles, concrete specimens exposed to alternate regimes had slightly higher relative compressive strength than that exposed to freezing–thawing-only cases, especially at the early testing ages. This indicated that alternate actions did not degrade the concrete continuously like freezing–thawing-only cases, and 7 days' carbonation would have a repairing effect on the strength loss of concrete after every 50 freezing–thawing cycles through carbonate deposition on its pores [13]. Therefore, carbonation improved the strength property of concrete [14, 15] degraded by freezing–thawing cycles. But it was worthy to note this effect reduced gradually with alternate repetitions, partially due to the narrowed pore structure in which carbonation filling exacerbates the freezing–thawing action.

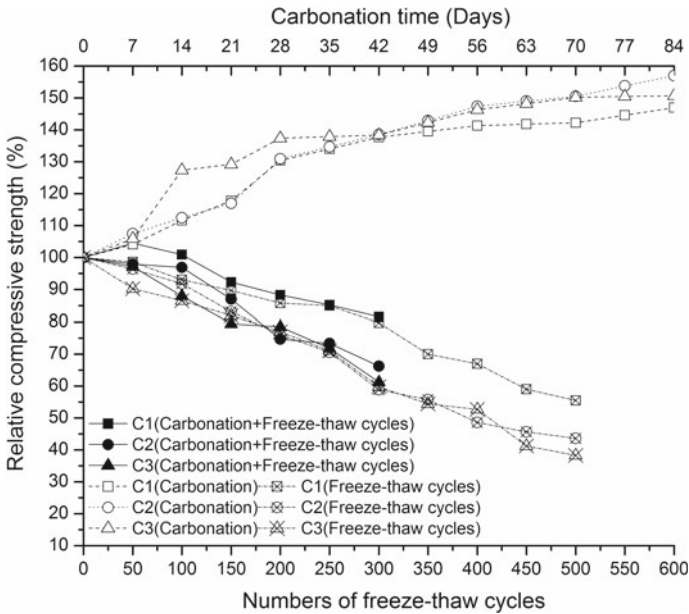


Fig. 1 Relative compressive strength of concrete exposed to different environmental actions

3.2 Relative Dynamic Elastic Modulus of Concrete Exposed to Different Environmental Actions

Figure 2 showed that the relative dynamic elastic modulus of concrete specimens exposed to carbonation-only cases increased with the carbonation time due to sufficient carbonation reaction. However, the relative dynamic elastic modulus of concrete specimens exposed to freezing–thawing-only cases or alternate carbonation and freezing–thawing actions decreased with the corresponding testing time. Under the same number of freezing–thawing cycles, the relative dynamic elastic modulus of concrete specimens exposed to alternate regimes was lower than that exposed to freezing–thawing-only cases. With the extension of alternate repetitions, the relative dynamic elastic modulus decreased sharply, especially for C3 specimens. This could be due to concrete shrinkage occurred when CO₂ gas penetrated into concrete and polymerized the silicate chains in C–S–H (which could decrease the volume, and consequently caused cracks) [16]. Building on the micro cracks caused by carbonation shrinkage, the freezing–thawing cycles contributed to their development, and yielded more microfractures [17]. Therefore, the alternate actions between carbonation and freezing–thawing cycles accelerated the frost damage of concrete, i.e. reduced the frost resistance of concrete.

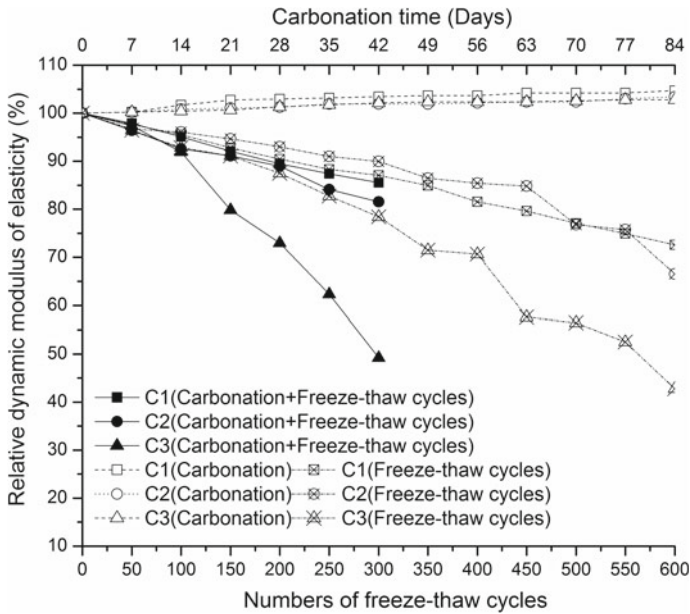


Fig. 2 Relative dynamic elastic modulus of concrete exposed to different environmental actions

3.3 Mass Change of Concrete Exposed to Different Environmental Actions

Figure 3 showed that the mass of concrete specimens exposed to carbonation-only cases increased slightly with the carbonation time. C2 specimens showed the largest mass increase, C3 specimens were the next, while C1 specimens were the least in mass increase. Some research results showed that the solid phase volume after carbonation increased 12–17% compared to its $\text{Ca}(\text{OH})_2$ volume, and led to a more compact concrete [18]. Moreover, the mass loss of concrete specimens exposed to freezing–thawing-only cases or alternate carbonation and freezing–thawing actions increased with the corresponding testing time. Under the same number of freezing–thawing cycles, the mass loss of concrete specimens exposed to alternate regimes was larger than that exposed to freezing–thawing-only cases. This indicated that the alternate actions between carbonation and freezing–thawing cycles accelerated the appearance damage of concrete.

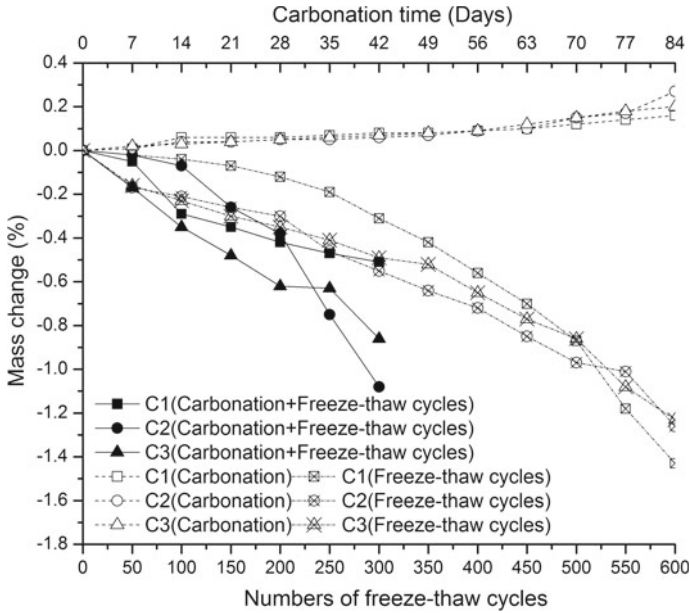


Fig. 3 Mass change of concrete exposed to different environmental actions

4 Damage Mechanics Fitting Model of Concrete Exposed to Alternate Carbonation and Freezing–Thawing Actions

According to the concept of macroscopic phenomenological damage mechanics [19], the damage variable of concrete could be expressed as follows:

$$D = \left(1 - \frac{E_d^i}{E_d^0} \right) \times 100 \tag{1}$$

where D is the damage variable of concrete after certain deterioration actions; E_d^i is the dynamic elastic modulus of concrete after these deterioration actions; E_d^0 is the dynamic elastic modulus of concrete before these deterioration actions.

4.1 Damage Mechanics Fitting Model of Concrete for Carbonation-Only Cases

The damage values of concrete exposed to carbonation-only cases were calculated via the above Eq. (1). Based on the damage data, a damage mechanics fitting model of concrete for carbonation-only cases was established as a logarithm expression,

Table 2 Parameter calibration values of the damage mechanics fitting model of concrete for carbonation-only cases

Code	a_c	b_c	R^2
C1	- 1.6175	2.5987	0.9761
C2	- 1.1346	2.2320	0.9068
C3	- 1.2072	2.5418	0.9396

and actually it was a negative damage.

$$D = a_c \ln(t) + b_c \quad (2)$$

where D is the damage values of concrete after t days' carbonation; t is the carbonation time; a_c and b_c are the fitting parameters, which could be determined via regression.

Table 2 showed that all of the R-square values of the fitting equations were 0.91 or more, which indicated an excellent correlation between carbonation time and concrete damage.

4.2 Damage Mechanics Fitting Model of Concrete for Freezing–Thawing-Only Cases

Similarly, the damage mechanics fitting model of concrete for freezing–thawing-only cases was established as a binomial expression.

$$D = a_f N^2 + b_f N + c_f \quad (3)$$

where D is the damage values of concrete after N freezing–thawing-only cycles; N is the number of freezing–thawing-only cycles; a_f , b_f and c_f are the fitting parameters, which could be determined via regression respectively.

Table 3 showed that all of the R-square values of the fitting equations were 0.97 or more, which indicated an excellent correlation between the number of freezing–thawing-only cycles and concrete damage.

Table 3 Parameter calibration values of the damage mechanics fitting model of concrete for freezing–thawing-only cases

Code	a_f	b_f	c_f	R^2
C1	0.000004	0.0428	0.3667	0.9977
C2	0.000090	- 0.0071	4.1861	0.9711
C3	0.000070	0.0499	0.5536	0.9881

4.3 Damage Mechanics Fitting Model of Concrete Exposed to Alternate Carbonation and Freezing–Thawing Actions

Considering the relationship between the damage values of concrete exposed to alternate regimes and those exposed to carbonation-only and freezing–thawing-only cases, a multivariable linear function could be fitted through the obtained data.

$$D_{c\&f} = k_c D_c + k_f D_f + k \tag{4}$$

where k_c is the carbonation impact factor; k_f is the freezing–thawing impact factor; k is the test constant; D_c , D_f and $D_{c\&f}$ are defined as below, with i stands for the time of alternate repetition and 0 stands for the original strength. The fitted parameters were shown in Table 4.

$$D_c = \begin{bmatrix} 1 - \frac{E_{d,c}^0}{E_{d,c}^0} \\ \dots \\ 1 - \frac{E_{d,c}^i}{E_{d,c}^0} \\ \dots \\ 1 - \frac{E_{d,c}^n}{E_{d,c}^0} \end{bmatrix}, D_f = \begin{bmatrix} 1 - \frac{E_{d,f}^0}{E_{d,f}^0} \\ \dots \\ 1 - \frac{E_{d,f}^i}{E_{d,f}^0} \\ \dots \\ 1 - \frac{E_{d,f}^n}{E_{d,f}^0} \end{bmatrix}, D_{c\&f} = \begin{bmatrix} 1 - \frac{E_{d,c\&f}^0}{E_{d,c\&f}^0} \\ \dots \\ 1 - \frac{E_{d,c\&f}^i}{E_{d,c\&f}^0} \\ \dots \\ 1 - \frac{E_{d,c\&f}^n}{E_{d,c\&f}^0} \end{bmatrix} \tag{5}$$

Table 4 showed all of the R-square values of this damage mechanics fitting model were 0.98 or more, which indicated an excellent correlation among the carbonation time, the number of freezing–thawing cycles and concrete damage under alternate regimes. The completed damage mechanics fitting model of concrete exposed to alternate carbonation and freezing–thawing actions could be expressed as:

$$\begin{cases} D_{c\&f} = 0 & t = 0, N = 0 \\ D_{c\&f} = k_c[a_c \ln(t) + b_c] + k_f[a_f N^2 + b_f N + c_f] + k & t > 0, N > 0 \end{cases} \tag{6}$$

The significant test of the above equation was conducted via the F test statistics. Table 5 showed this damage mechanics fitting model of concrete was significant, i.e. the development of concrete damage under alternate carbonation and

Table 4 Parameter calibration values of the damage mechanics fitting model of concrete for alternate carbonation and freezing–thawing actions

Code	k_c	k_f	k	R^2
C1	– 0.287	1.058	– 0.417	0.999
C2	– 3.833	1.090	– 0.765	0.986
C3	– 3.586	2.274	– 6.346	0.983

Table 5 Significant test results of the damage mechanics fitting model of concrete for alternate carbonation and freezing–thawing actions

Code	C1	C2	C3
F	1361.191	105.742	84.679
Sig	0.000	0.002	0.002

freezing–thawing cycles could be described via the data of concrete damage from the corresponding single environmental factors.

5 Conclusions

An experimental research on the influence of alternate carbonation and freezing–thawing actions on the mechanical properties and frost resistance of concrete was conducted. Through comparison with the experimental results obtained from carbonation-only and freezing–thawing-only cases, the following conclusions were reached:

- (1) The effect of carbonation on mechanical properties of concrete is complicated. Carbonation contributes to more compactness of concrete, whereas yields micro cracks at the surface of concrete due to carbonation shrinkage. This contributes strength increase at the early stage, but deteriorates quickly at later freezing–thawing cycles. Overall, the damage of concrete exposed to alternate carbonation and freezing–thawing actions is larger than that exposed to freezing–thawing-only cases.
- (2) The variation of relative dynamic elastic modulus and mass loss of concrete exposed to alternate carbonation and freezing–thawing actions is different with that of relative compressive strength. This indicates that these cracks or microfractures caused by carbonation are extremely small, which don't yield adverse effect on the compressive strength of concrete at early stage, but increase the frost damage and appearance damage of concrete by the expansion of cracks from the surface to interior.
- (3) A damage mechanics fitting model of concrete exposed to alternate carbonation and freezing–thawing actions is expressed as a multivariable linear function, in which two independent variables (carbonation time and number of freezing–thawing) are taken into consideration.

Acknowledgements This research was funded by the Key Research Plan of Shandong Province, China (2019GSF109057).

References

1. Wu ZW (1991) Concrete durability syndrome and its prevention and treatment. *Concrete*:4–8
2. Sun W, Zhang YM, Yan HD (1999) Damage and its resistance of concrete with different strength grades under double damage factors. *Cement Concr Compos* 21:439–442
3. Sun W (2006) Durability and service life of structure concrete under load and environment coupling effects. *J Southeast Univ (Nat Sci Edn)* 36:7–14
4. Ramezani-pour AA, Ghahari SA, Esmaeili M (2014) Effect of combined carbonation and chloride ion ingress by an accelerated test method on microscopic and mechanical properties of concrete. *Constr Build Mater* 58:138–146
5. Conciatori D, Sadouki H, Brühwiler E (2008) Capillary suction and diffusion model for chloride ingress into concrete. *Cem Concr Res*:1401–1408
6. Conciatori D, Laferrrière F, Brühwiler E (2010) Comprehensive modeling of chloride ion and water ingress into concrete considering thermal and carbonation state for real climate. *Cem Concr Res*:109–118
7. Qiao PZ, Chen FL (2013) Cohesive fracture and probabilistic damage analysis of freezing-thawing degradation of concrete. *Constr Build Mater* 47:879–887
8. Papadakis VG, Vayenas CG, Fardis MN (1991) Experimental investigation and mathematical modeling of the concrete carbonation problem. *Chem Eng Sci* 46:1333–1338
9. Xiao QH, Niu DT (2015) Damage model for concrete subjected to freeze-thaw cycles and carbonation. *J Build Mater* 18:763–766
10. Zhao GS, He Z, Yang HM (2013) Study of concrete durability under alternative effect of carbonization and freeze-thaw cycles. *Eng J Wuhan Univ* 46:604–609
11. Jiang LH, Na BB, Zhou XM, Wang T, Hu J, Xiong CS (2012) Tests for properties of concrete under alternate freeze-thaw and carbonation actions. *Adv Sci Technol Water Resour* 32:33–36
12. Glass GK, Page CL, Short NR (1991) Factors affecting the corrosion rate of steel in carbonated mortars. *Corros Sci* 32:1283–1294
13. Dias WPS (2000) Reduction of concrete sorptivity with age through carbonation. *Cem Concr Res* 30:1255–1261
14. Chang CF, Chen JW (2005) Strength and elastic modulus of carbonated concrete. *ACI Mater J* 102:315–321
15. Jerga J (2004) Physio-mechanical properties of carbonated concrete. *Constr Build Mater J* 18:645–652
16. Kropp TH, Hilsdorf HK (1989) The formation of silica gel during carbonation of cementitious systems containing slag cements. *ACI J* 114:1413–1428
17. Wang SR, Du X, Chen YL, Gao SS, Liang C, Ye PP (2015) Anti-frost property of concrete with different dosage of fly ash. *J. Univ Shanghai Sci Technol* 37:493–499
18. Fang J, Mei GX, Lu CR (1996) Influence of carbonation on properties of concrete. *Water Resour Hydropower Eng*:58–64
19. Li ZX (2002) *Damage mechanics and its application*. Science Press, Beijing

Research on the Top-Level Design of Urban Road Intelligent Promotion



Jin-dong Zhou

Abstract With the rapid development of the city, urban traffic congestion has become one of the main social problems. The development of urban intelligent transportation is an important solution. As an important carrier of intelligent transportation, urban road has gradually become the focus of urban development. By studying the development process of urban road intelligence, summarizing the development stages and achievements at home and abroad, combining with the latest Internet of things, cloud computing, big data, mobile Internet and other technologies, this paper puts forward the core development concept and overall framework of urban road intelligence promotion, and takes the middle high tech road of Zhangjiang Science City in Shanghai as an example to design the application scheme.

Keywords Urban road · Wisdom improvement · Top level design · Application scheme

1 Introduction

In May 2017, the national development and Reform Commission of the Ministry of housing and urban rural development issued the 13th five year plan for the construction of national urban municipal infrastructure [1], which proposed to promote the intelligent construction of municipal facilities and improve the level of safe operation management, including the improvement of urban road intelligence, the establishment of information interaction mechanism between road facilities and traffic subjects, and the improvement of traffic efficiency. According to the investigation of relevant literature, the intelligent improvement of urban municipal road proposed this time belongs to a new concept. Although the planning puts forward the development goal and vision, it does not explain the composition and Realization of the intelligent improvement of urban road in detail. Therefore, specific research is still

J. Zhou (✉)

Shanghai Pudong Architectural Design Research Institute Co., Ltd., Shanghai 201204, China

needed on how to carry out the intelligent improvement of urban municipal road to provide guidance for future planning and implementation.

2 An Overview of the Development Stage of Road Intelligence

2.1 Proposal of Intelligent Highway

Similar concept of road intelligent appeared in 1994. At the first world Intelligent Transportation Conference held in Paris, scholars and entrepreneurs from the West jointly discussed the current situation and future of intelligent transportation system. Among them, the intelligent highway has attracted great attention, and has gradually obtained huge funds for development in some countries. In 1995, the first International Conference on Intelligent roads and intelligent road exhibition were also held in Paris [2], France. Road traffic intelligence became an important development field of intelligent transportation.

2.2 Sustainable Development of Intelligent Highway

Combing the development of Intelligent Highway in foreign countries, we can see that the overall development of foreign countries is earlier, and now it has become a system, which has been explored and implemented in all aspects of the highway, and has preliminarily realized the intelligent highway management. For example, the United States first built an automated intelligent road in Southern California in 1997, the Smartway project in Japan in 2003, the smart in car project in Eindhoven, the Netherlands in 2006, and the E4 intelligent road in Sweden in 2008. By 2015, Michigan intelligent road was built [3], forming a continuous generation and continuous development of technology.

Compared with foreign countries, the intelligent development of domestic roads started nearly 10 years later, but the development after the 12th Five Year Plan is very rapid, and the construction of intelligent roads has been carried out in major cities in China. For example, in 2006, wankai intelligent highway was first built in Chongqing [4], and then intelligent highway construction was gradually promoted in Yunnan, Zhenjiang, Yangzhou, Anji, Haimen, Xuzhou, Changzhou and other cities. In addition, planning guidance documents at the national level have formed a relatively complete construction system.

Combined with the detailed cases of 19 smart roads at home and abroad, the main function of smart roads is to improve the “Safety + efficiency” of roads through information technology. The specific functions are summarized as follows (Fig. 1):

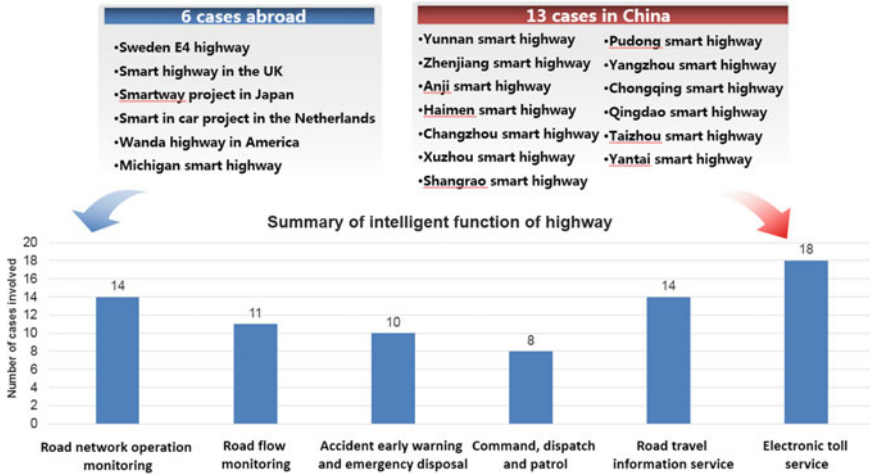


Fig. 1 Summary of intelligent function of highway

- (1) Road network operation monitoring
- (2) Road flow monitoring
- (3) Accident early warning and emergency disposal
- (4) Command, dispatch and patrol
- (5) Road travel information service
- (6) Electronic toll service

2.3 Organization Structure System of Highway Intelligence

The intelligent highway architecture consists of six core functions, which are distributed in 11 specific applications. The data is collected by the collection terminal and transmitted to the data analysis center by the wireless communication system. The analysis results are transmitted to the external through various release terminals, and finally serve the management department and driver users. The specific organizational structure system is shown in Fig. 2.

2.4 Proposal and Development of Urban Road Intelligent

The 13th five year plan of national urban municipal infrastructure construction issued by the national development and Reform Commission of the Ministry of housing

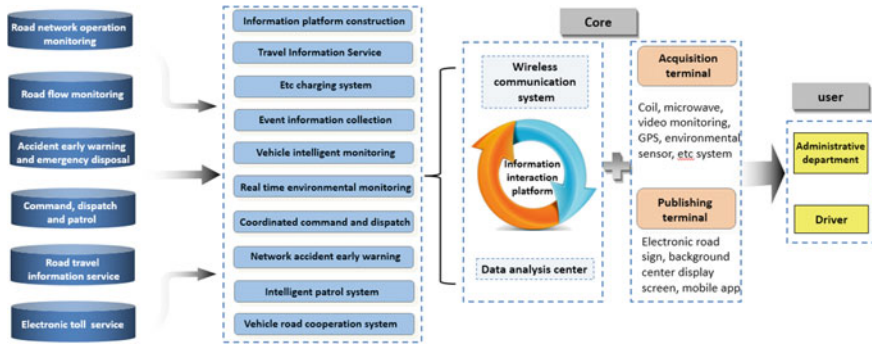


Fig. 2 Highway intelligent organization structure

and urban rural development clearly proposes to promote the intelligent construction of municipal facilities and improve the level of safe operation and management, including the improvement of road intelligence, the establishment of information interaction mechanism between road facilities and traffic subjects, and the improvement of traffic efficiency.

In China, Qiaoxiang road [5] in Shenzhen has been used as a specific application of road intelligent transformation. Qiaoxiang road is an important urban trunk road in the east–west direction of Shenzhen. The project crosses the two central districts of Futian and Nanshan, and undertakes the function of east–west traffic connection of Futian and Nanshan, and the traffic conversion function of South-North roads in Shenzhen. At present, the intelligent road reconstruction project has been implemented. With the help of the Internet of things, big data, artificial intelligence and other new generation information technology, the intelligent carrier of urban traffic information collection and release with data as the core has been constructed to realize the quality of road service, scientific management and efficient operation.

3 Top Level Design

3.1 Object Interpretation

3.1.1 Characteristics of Urban Road Development

The overall traffic environment of urban roads is complex, and the factors to be considered from the perspective of road intelligent transformation are more diverse. The main problems faced by urban roads are summarized as follows:

(1) Increase in public travel time poor travel timeliness

Many cities in our country are very congested during the rush hour of going to work. The problem of traffic jam has seriously increased the cost of economic transportation and affected the economic construction.

(2) Traffic congestion and frequent accidents

Traffic emergency response is the traffic management department in an emergency, through the emergency plan for emergency personnel and vehicles to ensure orderly rescue, reduce losses. The reasons for the slow response of traffic incidents are the slow information collection, slow decision-making and inadequate information release.

(3) Backward traffic management means—difficult law enforcement

Due to the backward traffic management means, a large number of police and manpower are needed to be deployed on the scene. It wastes a lot of manpower and financial resources to collect and process the traffic scene information.

(4) Environmental impact of road traffic—high noise emissions

The increase in the number of cars leads to another problem is the environmental impact. In the context of increasingly serious noise and air pollution, only comprehensive control and improvement from all levels of urban road traffic can reduce the impact on the environment.

Compared with highway, urban road has a big difference in design concept. The main service object of highway is motor vehicle, so the classification of highway mainly considers the size of traffic volume, while the technical index mainly considers the load and power of automobile, reflecting the concept of “vehicle oriented”; while urban road is located in densely populated and densely populated areas, its service objects include motor vehicles and non motorized vehicles Motor vehicles and pedestrians, especially people’s comfort and safety. The characteristic environment of highway design is closed, motor vehicle passing, no pedestrian participating and fast transportation, while the characteristic environment of urban road design is mixed traffic mode, composite function and urban environmental factors (Fig. 3).

3.1.2 The Core Development Concept

Urban road is composed of four elements: human, vehicle, road and environment.

- (1) Person: including traffic participants (drivers, pedestrians) and traffic managers (relevant departments, traffic police, etc.).
- (2) Vehicles: including motor vehicles (cars, buses, buses, trucks, special vehicles, etc.) and non motor vehicles (bicycles, electric vehicles, etc.).
- (3) Road: including roads (pavement, lane section, etc.), signs and markings and ancillary facilities (bus stops, street lamps, well covers, etc.)

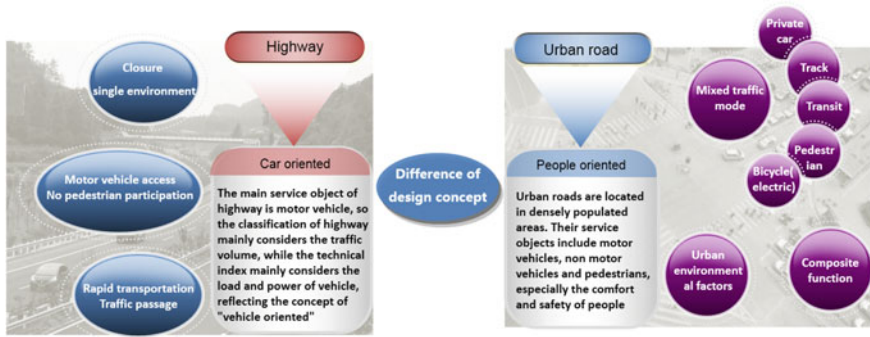


Fig. 3 The difference between highway and urban road

- (4) Environment: including hard environment (noise, climate, energy consumption, etc.) and soft environment (road operation management), etc.

The intelligent promotion of urban road should fully consider the actual needs of various elements, from how to form a system, how to be consistent with the development goals of the city and traffic, so as to achieve the coordination and overall improvement of various elements.

The core function framework of intelligent promotion of urban road is expanded and deduced on the basis of smart highway, and other functions are expanded based on the function of smart highway with strong correlation, combining with the characteristics and needs of urban road. The core of intelligent promotion of urban road is expanded from “security + efficiency” of highway to “security + efficiency + environmental protection”, fully considering the characteristics of various elements of urban road. The detailed deduction structure is as follows (Fig. 4):

3.2 Overall Framework

The overall framework of intelligent promotion of urban roads is composed of six modules. According to the logic relationship from the bottom to the top, they are collection facility module, cloud computing processing module, application module, publishing facility module and users, as well as the overall facility management and maintenance module (Fig. 5).

3.2.1 Acquisition Facility Module

The acquisition facility is the bottom layer of the whole architecture, and its main function is to collect peripheral basic data, which is the basis of system analysis. At present, the main acquisition devices include ground induction coil

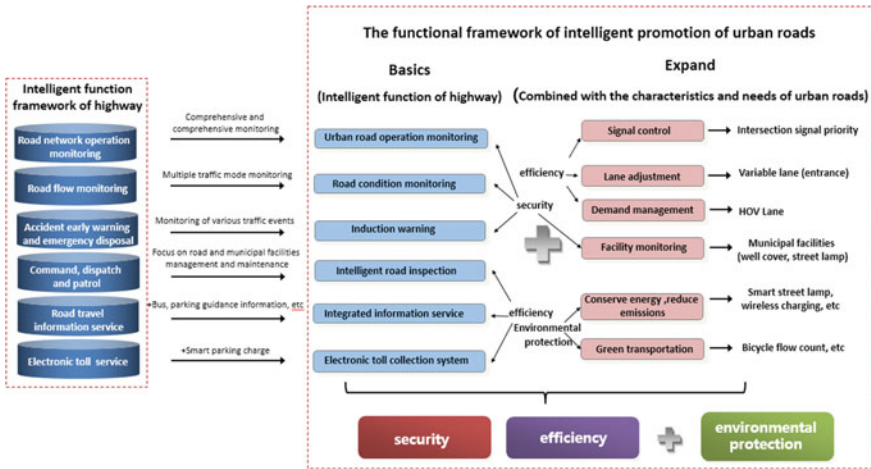


Fig. 4 The hierarchical deduction of the function of intelligent promotion of urban roads

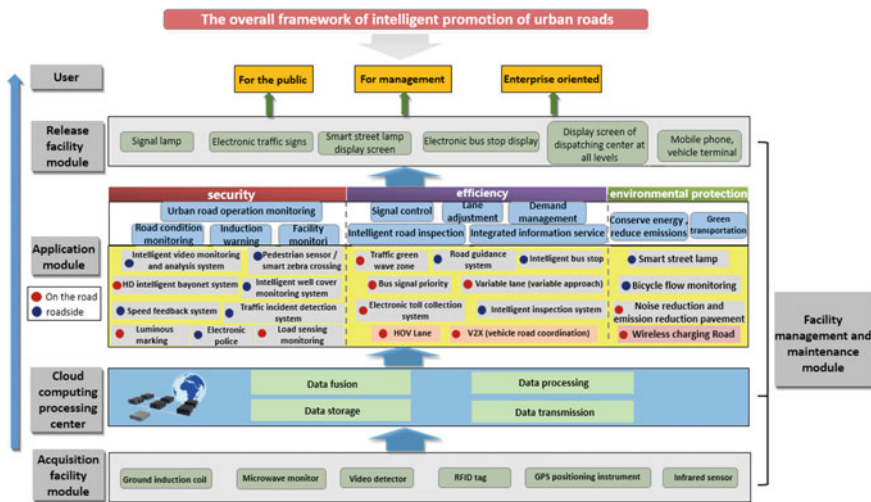


Fig. 5 Overall framework of intelligent improvement of urban roads

(flow measurement), micro blog detector (speed measurement), video detector (vehicle monitoring), RFID tag (identification), GPS (positioning), infrared sensor (detection).

3.2.2 Cloud Computing Processing Center

Build a road intelligent cloud computing processing platform, realize the collection of multi-department and multi-source data through various bottom collection facilities, access all kinds of original data sources and processing result data from external systems according to the transmission protocol, analyze and verify the received data, incorporate them into the database for preservation, and support all subsequent applications through big data cloud computing analysis. At the same time, as the core of background management, the whole platform can analyze the traffic flow on the road in real time, monitor the operation status of the pavement facilities in real time, optimize the road traffic organization in real time through big data analysis, and make the traffic management more intelligent.

3.2.3 Application Module

(1) Security layer

It is composed of four main functions, i.e. urban road operation monitoring, road condition monitoring, inductive early warning and facility monitoring. It is specifically composed of intelligent video monitoring and analysis system, electronic police, high-definition intelligent bayonet system, intelligent well cover monitoring system, speed feedback system, traffic event monitoring system, luminous marking and pedestrian inductive reminder.

(2) Efficiency layer

It is composed of five main functions: signal control, Lane adjustment, demand management, road intelligent inspection and comprehensive information service, specifically including traffic green belt, road guidance system, intelligent bus stop, bus signal priority, variable lane (variable entrance lane), electronic toll system, intelligent inspection system, HOV lane and v2x (vehicle road coordination), etc.

(3) Environmental protection

It is composed of two main functions: energy saving and emission reduction and traffic sharing, including intelligent street lamp, bicycle flow monitoring, noise reduction and emission reduction road and wireless charging road.

3.2.4 Release Facility Module

The release facility is the external information transmission medium in the whole architecture. Its main function is to display the analysis results to achieve the purpose of information transmission. At present, the main publishing devices include traffic signal lights, electronic traffic signs, intelligent street light displays, electronic bus stop signs display screens, dispatching center displays at all levels, mobile phones and vehicle mounted mobile terminals.

3.2.5 User

The service objects for the intelligent promotion of urban roads can be divided into three categories:

(1) Service delivery management department (to G)

It provides data and decision support for traffic management departments to alleviate traffic congestion, so as to improve traffic efficiency, provide information guidance for people's travel, and reduce environmental pollution.

(2) Serving the people (to C)

So that people can get fast traffic information, convenient arrangement of travel routes, so as to improve the ability to cope with traffic congestion and traffic events, reduce travel events and improve efficiency.

(3) Service to business (to B)

Through the data accumulation and openness of cloud platform, the data "incubation factory" can be formed for third-party enterprises to strengthen their functions and research and development.

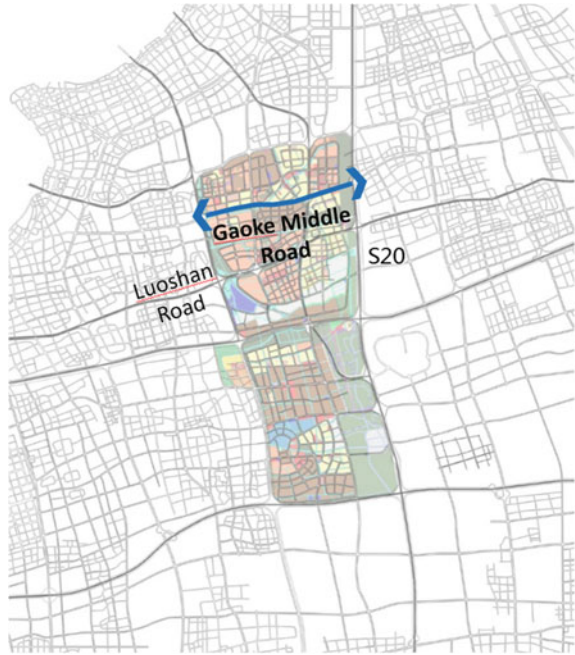
3.2.6 Facility Management and Maintenance Module

Facility management and maintenance mainly includes outfield equipment maintenance, communication network maintenance, central hardware maintenance and system software maintenance [6]. Through the layout of digital and intelligent infrastructure and related analysis technology, the track capture and tracking, information collection and push of using facilities can be realized, and the status quo and problems of facilities can be timely understood through data return to the platform, so as to shorten the response time for disposal to improve the management efficiency.

3.3 Main Application Scenarios

Starting from the core and framework of intelligent improvement of urban roads, high-speed roads and branch roads respectively correspond to special application directions. For example, high-speed roads are close to the concept of highways, without the interference of slow traffic, and relatively closed. They are urban rapid transportation systems; due to the small scale of the branch roads, such as signal control, Lane adjustment, etc., the space and necessity of implementation are not strong, so high-speed roads and high-speed roads are high-speed Roads and branches are used as general application scenarios. The primary and secondary trunk roads are the most important components of urban roads, and the degree of traffic mixing is the highest, which is most suitable for the scenarios to be served by the intelligent upgrading of urban roads, so they are taken as the key application scenarios.

Fig. 6 Location of Gaoke middle road



4 Application Scheme Example of Gaoke Middle Road

4.1 Necessity of Promotion

The section of Gaoke middle road (Luoshan road-s20) is located in the north area of Zhangjiang Science City. The function orientation of Zhangjiang Science City is to build an international first-class scientific research base, an important node of innovation resource allocation, an innovation system and mechanism exploration test area, and an innovation and entrepreneurship vitality science and technology city. Gaoke middle road is an important east–west trunk road in the north area of Zhangjiang Science City. The improvement of wisdom shows that the science city leads the development of science and technology (Fig. 6).

4.2 Promotion Principle

- (1) Combined with the construction and system of Gaoke middle road

This intelligent upgrading scheme fully combines the current situation of Gaoke middle road opinion facilities and systems to maximize compatibility, supplement and overall upgrade on the basis of the current situation.

- (2) Combined with background traffic flow and growth forecast of Gaoke middle road

From the perspective of future traffic growth, the corresponding intelligent promotion applications, such as variable lanes, can adapt to the development and change of future traffic in advance, especially considering the traffic increment brought by the gradual development of Zhangjiang Science City.

- (3) Combined with the nature of the surrounding land of Gaoke middle road

Since the transformation involves relevant service functions and facilities landing, the targeted transformation is carried out in full consideration of the surrounding land use nature, such as strengthening public transport service near residential areas and strengthening slow traffic around commercial areas.

- (4) Combined with the public transportation, slow traffic and static traffic of Gaoke middle road

Carry out a detailed investigation on the current situation of all kinds of traffic, and put forward a targeted transformation plan based on the current situation of public transport, slow traffic and static traffic.

4.3 Specific Plan

Combined with the above transformation principles, the intelligent transformation scheme of Gaoke middle road is as follows: 10 specific transformation applications + 1 prospect application + road management cloud platform construction, and the specific application is shown in Table 1 (Figs. 7 and 8).

5 Conclusions

Urban road construction is bound to develop in a more intensive direction, and traditional road infrastructure will be gradually replaced by intelligent facilities. With the progress of intelligent transportation technology, urban road construction will develop from people's perception of vehicles and roads to the interconnection of vehicles and roads. The intelligent improvement of urban road is a new concept and new direction. Relying on the information technology of the Internet of things, the collected data are processed and combined with the use of road resources to release the real-time information, so as to make the road operation and management more intelligent.

Table 1 Summary of specific application of intelligent promotion of Gaoke Middle Road

Serial No.	Application	Function	Facility layout
1	Intelligent video monitoring system	Real time monitoring, automatic vehicle search, accurate statistics of road traffic flow, provide evidence for punishment of violations, timely take early warning measures for emergencies, and reasonable scheduling	In the transformation of current monitoring probe, intelligent analysis elements (about 45 sets) are added, 3 high point monitoring points are set, and background intelligent video analysis system (including analysis and storage server) is added
2	Traffic signal intelligent control system	It realizes the real-time scheme selection control of the center, realizes the control mode of line coordination and induction coordination based on the system, and achieves the maximum traffic capacity of the road during the traffic peak period	Scats control is adopted in all signalized intersections
3	Road traffic intelligent guidance system	Realize the function of driving guidance, traffic information release and manual intervention	Display screens (including 4 nodes of inner ring, middle ring and outer ring) are arranged at main elevated entrances and intersections, and information transmission system (including server and switch) is added
4	Intelligent bus stop sign service system	Real time bus arrival forecast, instant information release, mobile intelligent monitoring, multimedia information service	There are 4 bus stops with relatively more bus lines: LED real-time arrival information display module (13 sets), information centralized control module (13 sets of components, 1 set of system), LCD query and advertising information LCD screen module (13 sets), background data management platform (access to the bus company)

(continued)

Table 1 (continued)

Serial No.	Application	Function	Facility layout
5	Intelligent well cover monitoring system	Equipment condition monitoring, precise navigation and positioning, remote monitoring, system data analysis, background alarm processing	At present, all manhole covers are installed with RFID tags (about 430), GPS / GPRS mobile receiving base stations (about 11 in 500 m receiving range), intelligent street lamp monitoring cameras (42 sets set at 300 m interval), and additional monitoring information background management system
6	Intelligent street lamp integrated management and control system	Intelligent lighting (energy saving), information release, video intelligent monitoring, WiFi hotspot	All the existing street lamps are transformed into intelligent dimming (about 430 intelligent dimming elements and 64 dimming receivers). The current streetlights are replaced with intelligent street lamp poles (with control box, about 40 lamps) at a spacing of 300 m. The corresponding modules are installed according to the needs. The display module is installed at important intersections and bus stops, and the monitoring module is installed in combination with the intelligent well cover system
7	Variable lane	Make full use of road resources to relieve directional traffic pressure	Scope: Jinke road to tanglu Road, reconstruction of ground marking (about 5.5 km), variable sign system gantry (8 sets), timing controller (8 sets)

(continued)

Table 1 (continued)

Serial No.	Application	Function	Facility layout
8	Luminous zebra crossing, pedestrian induction signal lamp	The self luminous zebra crossing line can store energy in the daytime and emit light at night to ensure the safety of pedestrians; the pedestrian crossing sensor will send out different voice prompts and warnings in different states of traffic lights to ensure the safety of crossing the street	Important trunk road intersections and pedestrian crossing places (Keyuan Road, Jinke road and Zhangdong Road intersection), luminous zebra crossing line (about 4 km), pedestrian crossing sensors (12 sets)
9	Speed feedback system	Speed real-time monitoring, prompt overspeed information, automatic speed measurement, intelligent display, not affected by the weather, traffic flow auxiliary statistics	Microwave speed measuring radar device can be set in combination with the current electronic police, and solar power supply mode is selected, and (6 locations) are set at the access viaduct
10	Bicycle flow counting system	Bicycle traffic counting, behavior data collection, understanding of travel habits, data analysis / release / sharing, promotion of green travel	Considering the surrounding residential concentrated travel area, bicycle flow counting display is installed at the section of non motorized lane at main intersection (4 places)
11	Vehicle road coordination (Prospect of unmanned driving)	Road safety service, automatic parking service, emergency vehicle yield, automatic car following service, unmanned driving (prototype)	The running vehicles are equipped with RFID, GPS, vehicle and roadside camera monitoring system, D2D information transmission and sharing system, and roadside network transmission base station



Fig. 7 Design sketch of intelligent promotion of Gaoke middle road—Road section



Fig. 8 Design sketch of intelligent promotion of Gaoke middle road—Intersection

References

1. The 13th five year plan of municipal infrastructure construction in China: Ministry of housing and urban rural development, National Development and Reform 2017
2. Sun W (1995) Intelligent highway traffic in the 21st century. World Sci 1995(6)
3. Xu Z, Li J (2019) A review on Intelligent road and its related key technologies. Zhongguo Gonglu Xuebao/China J Highway Transp 32(8):1–24
4. Biao H (2007) Application of automatic video traffic accident detection system on Chongqing wankai Expressway. China J
5. Zhou O, Xin Z, Wei X (2019) Construction and thinking of urban smart road. In: China urban transport planning annual meeting in 2019
6. Na Z (2017) Analysis and discussion on the implementation of the concept of smart city in the management and maintenance of municipal facilities and the realization of information management. Build Mater Constr Decor 008:15–16

Jin-dong Zhou (1989–) male, master, engineer, main research direction: urban traffic planning and traffic design

Predicting Red Light Running Violation Using Machine Learning Classifiers



Muhammad Zahid, Arshad Jamal, Yangzhou Chen, Tufail Ahmed, and Muhammad Ijaz

Abstract Red light Running (RLR) violation remains as an important road safety concern at urban intersections. Existing method have mostly used statistical regression-based methods to explore factors contributing to RLR. However, it is well-known that statistical methods are based on predefined associations among variables and are unable to capture latent heterogeneity. This study aims to classify and predict RLR using spatial analysis and machine learning (ML) methods. Georeferenced RLR violation data for the year 2016 was collected for the city of Luzhou, China. To identify violation hotpots, frequency-based clustering was carried out using collect event tool in ArcMap geographic information system (GIS). Prior to RLR prediction via ML, data imbalance problem was addressed using a random over-sampling technique. Two widely used ML algorithms, i.e., Random Forest (RF) and Gradient Boosted Decision Tree (GBDT), were then used for prediction and classification of RLR. The performance of these models was assessed with accuracy and Cohen's kappa. The results showed that GBDT had an overall accuracy of 96% outperformed the RF.

M. Zahid

College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China
e-mail: zahid@emails.bjut.edu.cn

A. Jamal

Department of Civil and Environmental Engineering, King Fahd University of Petroleum and Minerals, KFUPM, Box 5055, Dhahran 31261, Saudi Arabia
e-mail: arshad.jamal@kfupm.edu.sa

Y. Chen (✉)

College of Artificial Intelligence and Automation, Beijing University of Technology, Beijing 100124, China
e-mail: yzchen@bjut.edu.cn

T. Ahmed

UHasselt, Transportation Research Institute (IMOB), Agoralaan, 3590 Diepenbeek, Belgium
e-mail: tufail.ahmed@uhasselt.be

M. Ijaz

School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China

Keywords Aggressive driving · Hotspot analysis · Machine learning · Red light running

1 Introduction

Red-light running (RLR) has been recognized as a major problem for traffic safety that can result in frequent traffic conflicts and also increases the severity of crashes. Statistics demonstrate that large number of crashes occur at signalized intersections due to traffic violations, specifically red-light running (RLR). In 2012, more than 2.5 million intersection related crashes occurred resulted in 2,850 fatal crashes and 680,000 injury crashes [1]. RLR violations have also been identified as a serious problem causing intersection associated crashes. The Insurance Institute for Highway Safety identified that in 2012 the violation of RLR resulted in 683 deaths and an estimated 133,000 injuries in the US [2]. Particularly, according to data collected in China from January 2012 to October 2012, RLR violation resulted in 4,227 serious injury crashes and 789 fatalities [3].

Statistics showed that a higher number of crashes happen at signaled intersections due to traffic violations, mainly violations of RLR [4, 5]. Indicators of RLR violation typically involve under the age of 30 years old, with poor driving records, getting closer to the vehicle in front, not wearing a safety belt, feeling angry with other drivers, and use of alcohol [6]. McCartt and Wen [7] found that there was a statistically significant reduction in the number of violations occurring at 0.5 s (39%) and 1.5 s (86%) after the onset of the red signal [7]. However, few studies have considered RLR violations and the effects of penalties (e.g., fines, demerit points) used to sanction those programs [8–10]. With regard to the traffic conditions, wider intersections, and higher volumes of traffic were related to a higher rate of RLR violations [11]. The longer times are more likely to promote violation of RLR at intersections. The smaller cycle's length ensures more drivers' possibilities per hour of running a red light at each shift [12]. Lu et al. [13] implemented a randomized experiment in China and showed that informing drivers that they were observed committing traffic violations by automatic detection devices deterred drivers from committing the same traffic violation in the future.

Zhang et al. [14] proposed a probabilistic model for predicting RLR violations, considering the minimization of missing errors and false alarm rates [14]. A decision tree model was introduced to predict RLR violations. The results indicate that the distance of the vehicle at the yellow onset, driving speed at the yellow onset, and location in traffic flow were determined to be the most significant variables [15]. Support Vector Machine (SVM) and Hidden Markov Model (HMM) were utilized to predict violations of the RLR. The result showed that attributes such as distance-to-intersection (DTI), velocity, speed, time to an intersection (TTI), and the most useful attributes had been found to be the required deceleration parameter (RDP) [16]. Jahangiri et al. applied a random (RF) forest and SVM to predict RLR Violation. The result indicated that the monitoring period and critical instant time provides

necessary information for predicting RLR violations [17]. Moreover, a study was conducted to address the ability of artificial neural networks (ANN) to estimate the RLR cycle and predict the occurrence of RLR based on the four-status vehicles (DTI, pace, headway, and a number of front vehicles) at the yellow start [18]. Zahid et al. used a stacked generalization method to predict and classify the aggressive driving behavior of taxi drivers. The analysis also addressed different types of violations, including overspeeding, wrong-way driving, prohibited lane marking, and RLR [19].

Existing research has aimed at analyzing and modeling the characteristics of red-light violation at signalized intersections with statistical methods. Only a few research studies have attempted to predict and classify RLR violations with machine learning algorithms taking into account different vehicles types. In this study, we focused on RLR violations prediction and classification committed by different vehicles for a case study in China. The main contributions of the study are (i) developing frequency-based hotspots maps to identify RLR prone and high-risk zones in the study area, (ii) adopting machine learning algorithms such as GBDT and RF to predict and classify the RLR violations committed by aggressive driving. (iii) evaluating the predictive performance of proposed methods using K-fold cross-validation.

The remainder of the paper is organized as follows. Section 2 presents the study area. Section 3 provides the descriptive statistics analysis for collected data. Section 4 presents RLR hotspots methods and analysis. Section 5 presents the methods adopted machine learning for RLR prediction. Section 6 provides results and discussion. Finally, Sect. 7 summarizes key study findings and outlook for future studies.

2 Study Area

Luzhou; is a prefecture-level municipality with an area of 101.05 km², and a population over 1 million, located in the southeast of Sichuan Province, China. Located at the combination of the Tuo River and Yangtze River, the Luzhou port on the Yangtze River is not only the major port of Sichuan since the Chongqing Province in 1997. As per the National Bureau of the Statistics People's Republic of China (PRC), by there of 2017, the country had 4.77 million paved roads and over 300 million registered vehicles [20]. The study area can be seen in Fig. 1.

3 Data Description

RLR data for the year 2016 was taken from the Sichuan traffic police department. The data was extracted from off-site traffic monitoring and enforcement cameras that are installed at selected important locations. Georeferenced violation data was collected on various attributes including detailed information on latitudes and longitudes, temporal attributes (such as peak/off-peak periods, weekdays/weekends, the

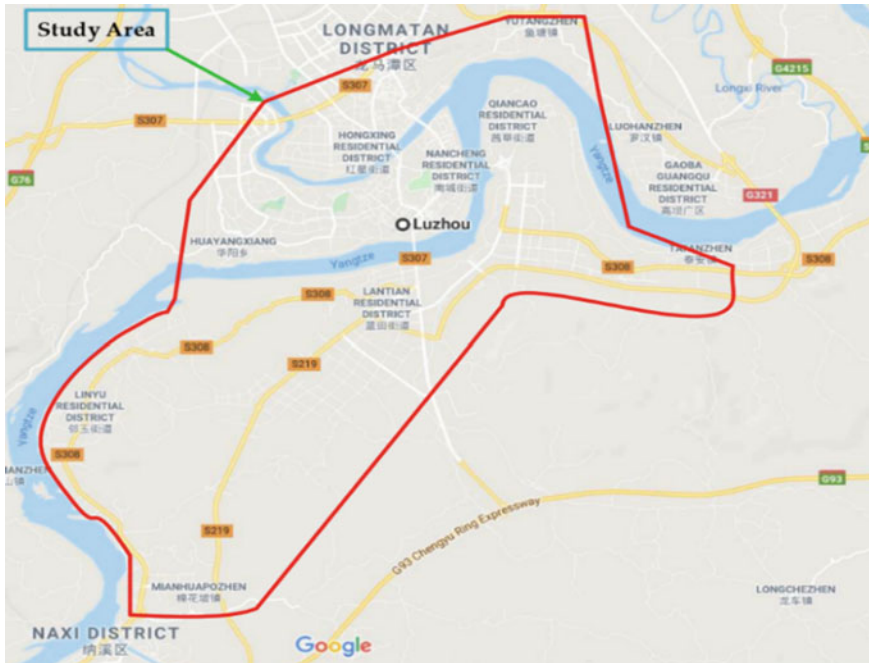


Fig. 1 Study area (from google map)

season of the year), road types (general urban road) vehicle types (private car, taxi, van, small truck, and bus), and resulting RLR violation. Latitude and longitude attributes were used to approximate the location of violation occurrence. The RLR violation committed by private car (around 66%) was the highest, and that those by buses were the lowest (approximately 3%). Similarly, winter had the highest proportion (38.43%) of observed violations followed by autumn (23.20%), spring (23.11%), and summer (15.25%). Moreover, the total count of RLR violation committed in peak hours and off-peak hours are 32.57 and 67.43. Likewise, the percentage of the total RLR violation occurred in weekday and weekend were 68.60 and 31.40%, respectively. The detailed descriptive statistic of RLR violation are given in Table 1.

4 RLR Hotspot Analysis

The distribution of hotspots was identified using Spatial statistical analysis in ArcMap. Hotspot analysis was carried out in four different stages. In the first step, red-light traffic violation point data was added to GIS. Data clusters were created to convert the data to weighted points in the second step. Collect Event (spatial statistics tool) was used to convert the red-light traffic violation data to weighted points,

Table 1 Descriptive statistics of RLR violation (N = 3,258)

Factor	Percentage of total RLR violation (%)	Frequency
<i>Vehicle type</i>		
Private car	66.30	2160
Taxi	19.71	642
Van	6.17	201
Bus	5.22	170
Small truck	2.61	85
<i>Seasons</i>		
Winter	38.43	1252
Autumn	23.20	756
Spring	23.11	753
Summer	15.25	497
<i>Nature of Weekday</i>		
Weekdays	68.60	2235
Weekends	31.40	1023
<i>Hours of the day</i>		
Peak hours (9:00 am–11:00 am, 15:00 pm–17:00 pm)	32.57	2197
Off peak hours (11:00 am–15:00 pm, 17:00 pm–9:00 am)	67.43	1061

and the weights to the data are assigned based on its frequency. Collect Event tools combines all coincident points that have the same X and Y centroid coordinates. Further, to identify the hotspots in the study area, Getis-Ord statistic was used on weighted data. A value close to zero shows a random distribution in the observed. Mathematically, Getis-Ord statistic, and its z-score are expressed in terms of Eqs. (1) and (2), respectively.

$$G_i = \frac{\sum_{j=1}^n w_{ij}x_j}{\sum_{j=1}^n x_j} \tag{1}$$

$$Z(G_i) = \frac{\sum_{j=1}^n w_{ij}x_j - \bar{x} \sum_{j=1}^n w_{ij}^2}{\sqrt{VAR(G_i)}} \tag{2}$$

where G_i indicate the spatial dependency of the incident i overall n events, x_j is the magnitude of variable X at incident location j overall nor the severity index at location j , w_{ij} represents the weight value between target neighbor pair (i and j that represents their spatial interrelationship), and \bar{x} is the sample mean.

High value of Getis-Ord statistic means that a cluster has high index values which indicate a hot spot. Similarly, a low value of Getis-Ord statistic shows a cluster of

low index values representing a cold spot. In the final step, the Inverse Distance Weighted (IDW) method of interpolation was applied to estimates red light violation hotspot in the study area, which is shown in Fig. 2. In IDW, the neighboring values are estimated by averaging the values of sample data points. The closer a point is to the center of the cell being estimated, the more influence or weight; it has in the averaging process.

A total of 3,258 traffic violations were observed in the study area. Based on the analysis, there are three hotspots observed in the study area. Two hotspots are concentrated in a residential area in the west of the study area. Both of these hotspots are very close to each other. While the third hotspot is located in the eastern part of the study area near to the commercial area, it is worth explaining from Fig. 2 that all the three hotspots out of the city with decreased built-up areas. The traffic is usually

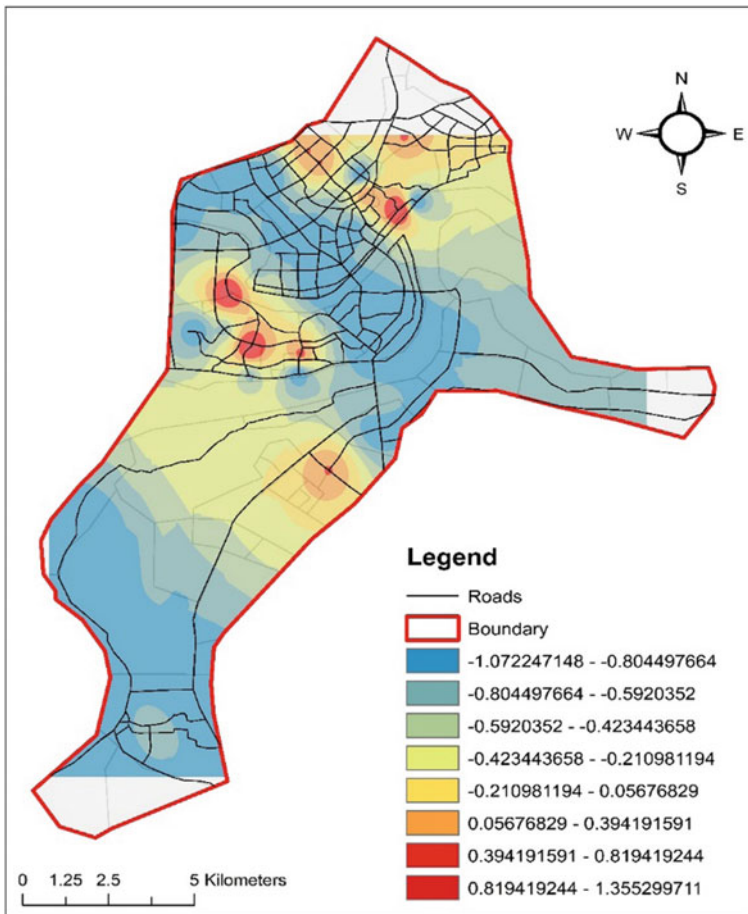


Fig. 2 Distribution of RLR hotspots in the study area

monitored strictly in the central business districts (CBD) traffic is through traffic surveillance cameras at the traffic signals, which discourage red-light violations in CBD. Moreover, the number of red-light traffic violations on low traffic volume roads in low-density residential area approaches to be higher.

5 Prediction of RLR Using Machine Learning

The raw data contain some classes with fewer instances that create an imbalanced learning problem. Machine learning algorithms mostly find it hard to manage an imbalanced dataset. For imbalanced datasets, the model's learning ability is weak. Since our objective is to predict as accurately as possible positive instances. There are several approaches for handling data with imbalanced distributions. Research has provided a solution to sampling imbalance problem, and a detailed description can be found in [21]. Under-sampling and oversampling are techniques in data processing used to modify the distribution of the classes. Meanwhile, the under-sampling method frequently results in the loss of the major sample information. In this study, the technique of random oversampling was used to create artificial samples. Random oversampling technique introduces a set that is obtained randomly from minority class. The aim of achieving a balanced degree is to reproduce the instances and to increase the total number of minority instances.

5.1 *Random Forest (RF)*

Random Forests (RF) is an evolution of Breiman's bagging concept (2001) and has been developed to boost competitors [22]. RF's principle is to incorporate a set of decision trees that are formed by bootstrapping the sample and randomly choosing a subset of inputs at each node. (1) The algorithm randomly chooses a given number of predictors at each node and determines the best split within that subset only, and (2) RF has no pruning step. In RF, the standard process is applied to calculate variable importance for each tree. The first step is to transfer data from out-of-bag (OOB) down the tree and determine the predicted values. The OOB sample here is a collection of data that will not be used to create the current tree. Lastly, in the OOB data, the variable's values are randomly permuted, while other independent variables are fixed. The configured OOB sample runs down the tree and yet again computes the predicted values. Lastly, the variable value significance is computed by averaging the difference in the error rate for all trees before and after the permutation. To implement an RF model, two parameters should be adjusted: (1) the number of trees to grow in the forest, and (2) the number of layers to split the tree nodes. Moreover, selecting a higher value of layers helps increase the correlation between trees that yield identical results while voting. The general formula of the G_{ini} index is defined in Eq. (3).

$$Gini\ Impurity(p) = 1 - \sum_{i=1}^N p_i^2 \quad (3)$$

where p is the dataset, N is the number of classes, p_i is the i class frequency in p the dataset.

5.2 Gradient Boosted Decision Trees (GBDT)

GBDT is an ensemble learning algorithm. GBDT is based on Boosting, and the classification result of this algorithm is the accumulation of the result of each tree. Therefore, GBDT is more concerned about reducing the deviation during training. In addition, it has some advantages, including the ability to find non-linear transformations and handle skewed variables without requiring transformations. Also, GBDT has good robustness and high scalability. Furthermore, GBDT sequentially combines weak learners so that each new learner fits the residue from the previous step to strengthen and improve the model. The final model aggregates the outcomes of each step and achieves a good learner. The accuracy of the model depends on the estimator and learning rate. The estimator is the number of trees used in the model. Learning rate simply means how fast the models learn. Every added tree adjusts the overall model, and the magnitude of the adjustment tends to depend on the learning rate. If the learning rate is less, the more tree will be needed to train the model [23].

6 Model's Evaluation Metrics

In this study, we used the most common assessment indicators to check the performance of the different classification models. Accuracy is the proportion of the correct sample to the total number of samples and can be calculated from the Eq. (4). Similarly, Cohen's kappa is the probability of agreement take away the probability of random agreement divided by one minus the probability of random agreement, which can be calculated from Eq. (5).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$k = \frac{p_o - p_e}{1 - p_e} \quad (5)$$

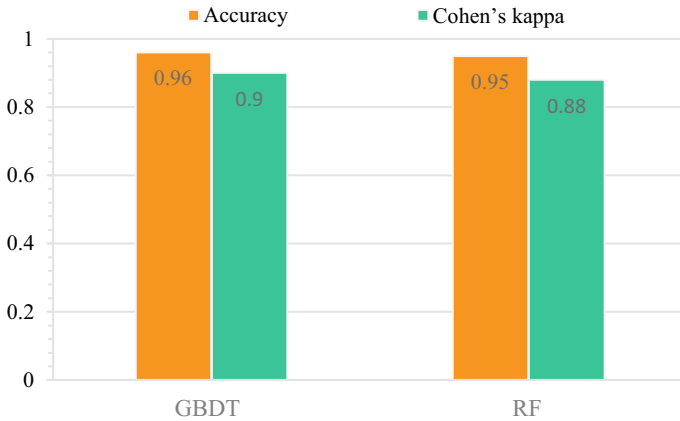


Fig. 3 Model Comparison expressed in accuracy and Cohen's kappa

7 Results and Discussions

7.1 Model's Comparison

The model comparison is made to assess the efficiency of the applied models. The performance of the model is evaluated with regard to the accuracy, Cohen's kappa. The parameters include in GBDT were tree depth or n estimators, Numbers of models, and learning rate. The accuracy and Cohen's kappa achieved when the values of these parameters were 10, 100, and 0.1. The accuracy and Cohen's kappa achieved for GBDT is 0.96 and 0.95, as shown in Fig. 3. Similarly, On the other hand, the parameters included in the random forest were tree depth, the minimum number of nodes, and the number of models. The accuracy and Cohen's kappa achieved when the values of parameters were 10, 1, and 100. The accuracy and Cohen's kappa value achieved for the random forest is 0.90 and 0.88. Figures 4 and 5 shows the stratified K-Fold error rate for random forest and GBDT. The results for these models were obtained using stratified 10-Fold cross-validation.

7.2 Spearman Correlation Analysis

We estimate the rank correlation coefficient of the Spearman between the two features and obtain the matrix of the correlation coefficient. It aims to assess how well a monotonic function could be used to represent the relation between two variables. The correlation matrix describes the correlation and no correlation variance of a range of features through the values created in matrix ranging from + 1 to - 1. + 1 means high and positive correlated while - 1 means less and negative correlated.

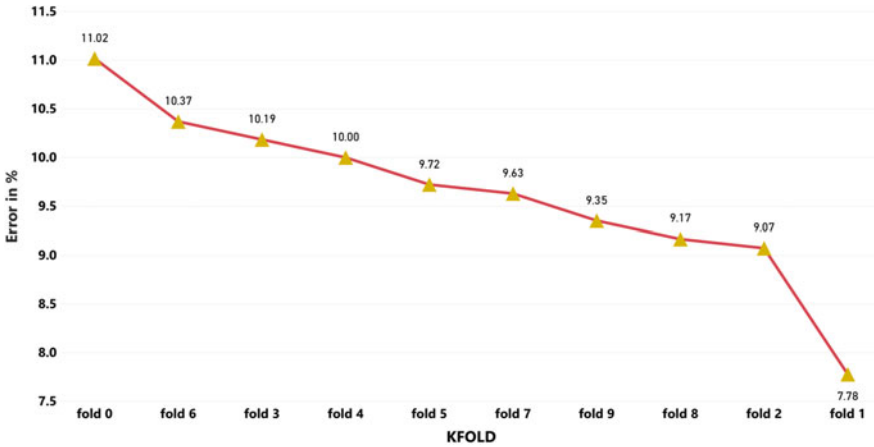


Fig. 4 K-fold error rate for RF (expressed in %)

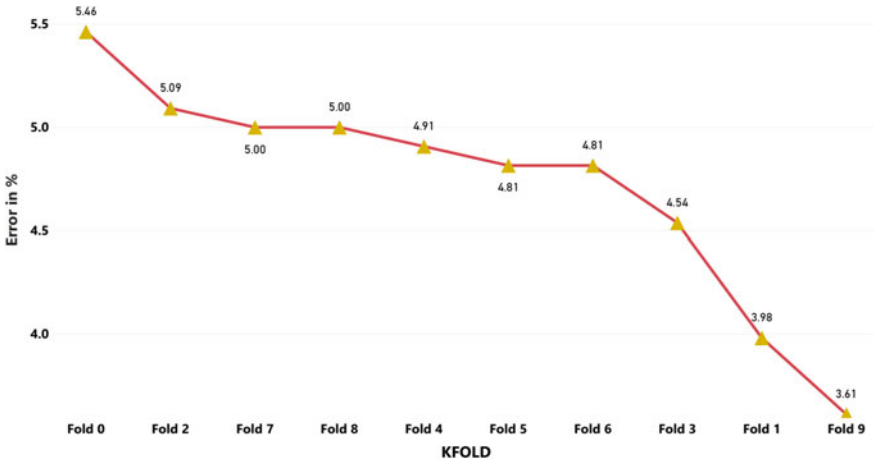


Fig. 5 K-fold error rate for GBDT (expressed in %)

The features which are less and negative correlated are season and date or date and season as shown in Fig. 6.

8 Conclusions

This study utilized various machine learning algorithms for prediction and classification of vehicles RLR violations data collected during 2016, in the city of Luzhou, China. First the raw data descriptive statistics showed that private cars

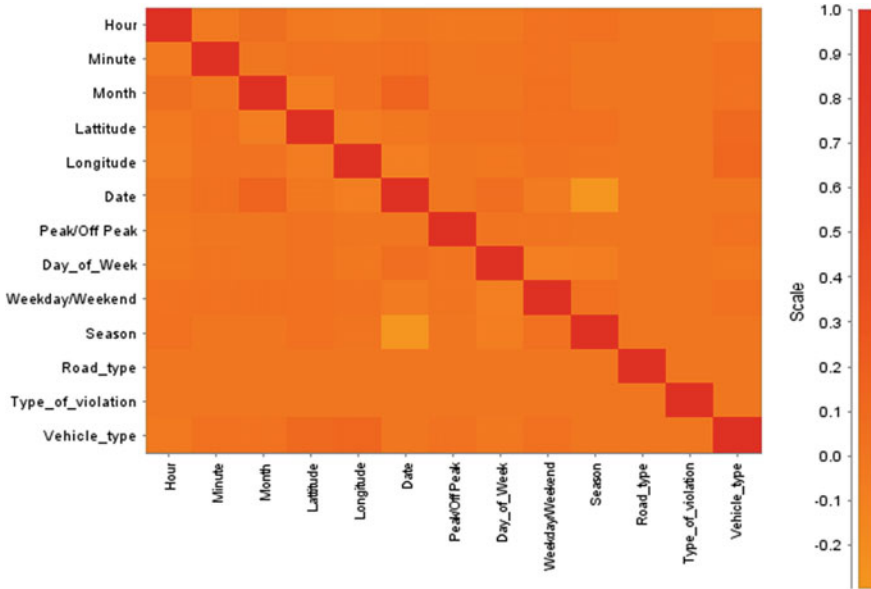


Fig. 6 Spearman's correlation matrix

(66.30%) committed most RLR violations followed by taxi (19.71%), van (6.17%), bus (5.22%), and small truck (2.61%). Frequency-based cluster analysis was also conducted to identify RLR hotspots. Accurate identification of such high-risk zones is vital for prioritizing sites where the effects of safety improvement strategies could be more promising. For RLR, prediction, raw data was preprocessed due to imbalance data problems with random over-sampling. Study results demonstrated that over-sampling significantly improved the predictive performance for both RF and GBDT methods. Both the models were found in predicting RLR violations; however, GBDT with accuracy 96% outperformed the RF. The study has some limitations which need to be addressed. The current data lack information on certain important factors (i.e., drivers socio-demographic characteristics of drivers, car license plates, local urban environment features, etc.) could be considered in future studies.

Acknowledgements The Authors acknowledge Sichuan Traffic Police Department for providing the violation data used in this study.

References

1. Administration NHTS (2012) National center for statistics and analysis, traffic safety facts 2009. US Department Transportation, Washington, DC

2. Gazis D, Herman R, Maradudin A (1960) The problem of the amber signal light in traffic flow. *Oper Res* 8(1):112–132
3. Wang X, Yu R, Zhong C (2016) A field investigation of red-light-running in Shanghai, China. *Transp Res Part F Traffic Psychol Behav* 37:144–153
4. Zahid M, Chen Y, Jamal A (2020) Freeway short-term travel speed prediction based on data collection time-horizons : a fast forest quantile regression approach. *Sustainability* 12(2):646
5. Jamal A, Rahman MT, Al-Ahmadi HM, Mansoor U (2020) The dilemma of road safety in the eastern province of Saudi Arabia: consequences and prevention strategies. *Int J Environ Res Public Health* 17(1):157
6. Retting RA, Williams AF (1996) Characteristics of red light violators: results of a field investigation. *J Safety Res* 27(1):9–15
7. McCart AT, Hu W (2014) Effects of red light camera enforcement on red light violations in Arlington County, Virginia. *J Safety Res* 48:57–62
8. Jamal A, Rahman MT, Al-Ahmadi HM, Ullah IM, Zahid M (2020) Intelligent intersection control for delay optimization: using meta-heuristic search algorithms. *Sustainability* 12(5):1896
9. Pulugurtha SS, Otturu R (2014) Effectiveness of red light running camera enforcement program in reducing crashes: evaluation using ‘before the installation’, ‘after the installation’, and ‘after the termination’ data. *Accid Anal Prev* 64:9–17
10. Porter BE, Johnson KL, Bland JF (2013) Turning off the cameras: red light running characteristics and rates after photo enforcement legislation expired. *Accid Anal Prev* 50:1104–1111
11. Porter BE, England KJ (2000) Predicting red-light running behavior: a traffic safety study in three urban settings. *J Safety Res* 31(1):1–8
12. Bonneson JA, Zimmerman K, Brewer MA (2002) Engineering countermeasures to reduce red-light-running. Texas Transportation Institute, Texas A&M University System College
13. Ou G, Gao Y, Liu Y (2012) Real-time vehicular traffic violation detection in traffic monitoring stream. In: 2012 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology, vol 3, pp 15–19
14. Zhang L, Zhou K, Zhang W, Misener JA (2009) Prediction of red light running based on statistics of discrete point sensors. *Transp Res Rec* 2128(1):132–142
15. Elmitiny N, Yan X, Radwan E, Russo C, Nashar D (2010) Classification analysis of driver’s stop/go decision and red-light running violation. *Accid Anal Prev* 42(1):101–111
16. Aoude GS, Desaraju VR, Stephens LH, How JP (2012) Driver behavior classification at intersections and validation on large naturalistic data set. *IEEE Trans Intell Transp Syst* 13(2):724–736
17. Jahangiri A, Rakha HA, Dingus TA (2015) Adopting machine learning methods to predict red-light running violations. In: 2015 IEEE 18th international conference on intelligent transportation systems, pp 650–655
18. Li P, Li Y, Guo X (2014) A red-light running prevention system based on artificial neural network and vehicle trajectory data. *Comput Intell Neurosci* 2014
19. Zahid M, Chen Y, Khan S, Jamal A, Ijaz M, Ahmed T (2020) Predicting risky and aggressive driving behavior among taxi drivers: do spatio-temporal attributes matter? *Int J Environ Res Public Health* 17(11):3937
20. Yearbook CS (2012) National Bureau of statistics of China. *China Stat Yearb*
21. He H, Garcia EA (2009) Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 21(9):1263–1284
22. Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
23. Wang Y, Feng D, Li D, Chen X, Zhao Y, Niu X (2016) A mobile recommendation system based on logistic regression and gradient boosting decision trees. In: 2016 international joint conference on neural networks (IJCNN), pp 1896–1902

Bi-Objective Subway Timetable Optimization Considering Changing Train Quality Based on Passenger Flow Data



Hanlei Wang, Peng Wu, Yuan Yao, and Xingxuan Zhuo

Abstract As an effective transport mode, rail systems have been developing rapidly in recent years. A reasonable timetable not only enables passengers to have a better travel experience, but also saves the operating cost of subway companies. Therefore, the timetable design is the key to the reliability of subway services. Based on actual passenger flow, this paper develops a new bi-objective integer programming model with the objectives of minimizing the total passenger waiting time and the net energy consumption of trains simultaneously. To effectively solve the proposed model, a problem-specific non-dominated sorting genetic algorithm is proposed. A case study based on Metro Line 1 of Fuzhou, China, illustrates the effectiveness and efficiency of the proposed approach.

Keywords Timetable · Energy · Passenger waiting time · Multi-objective optimization

1 Introduction

Subway system has developed rapidly in many cities around the world. For example, many cities in China, such as Beijing, Tianjin, Hong Kong, Shanghai, Guangzhou, Shenzhen and other cities have sound subway systems, and subways in many cities are also under construction. Owing to the comfort, speed and convenience, subway has become an important transportation means for people's daily mobility, and it has also become an important part of transportation in many cities [1]. To ensure the efficiency and effectiveness of the subway system operation, a reasonable subway schedule is essential. From a strategic and tactical perspective, train timetabling problem (TTP) is a key issue, which lays the foundation for the quality of service of the subway operating system [2]. Timetable optimization not only considers the travel experience of passengers, but also considers the operating costs of subway companies, and even the impact on the environment. The influence of many factors

H. Wang · P. Wu (✉) · Y. Yao · X. Zhuo

School of Economics and Management, University of Fuzhou, Fuzhou 350116, P. R. China

has caused the complexity of subway systems, so timetable optimization problems have attracted much attention from both scholars and practitioners.

The optimization of subway timetable plays a vital role in improving the efficiency of subway system operation. In recent years, many scholars have devoted themselves to the research of subway timetable optimization. They have studied the subway timetable in different ways from different perspectives. When considering the optimization of the subway system, it is necessary to comprehensively consider the revenue of subway companies, the travel experience of passengers, environmental pollution, and even certain social responsibilities. Therefore, the optimization of subway timetable is usually a multi-objective optimization problem. Yang et al. [3] formulate a two-objective integer programming model to improve utilization of renewable energy consumption and reduce passenger waiting time. Sun et al. [4] develops a bi-objective timetable optimization model to the pure energy consumption and minimize the total passenger waiting time. Yang et al. [5] formulate two-stage stochastic optimization model to optimize timetable and train speed curve to reduce train energy consumption. In order to balance energy use and passenger travel time, Yang et al. [6] design an optimized model of integrated energy distribution and passenger distribution. Zhou et al. [7] focuses on the problem of multi-periodic train timetabling, they construct a mixed integer programming model of one-way and two-way programming to minimize the total travel time. Abdolmaleki et al. [8] aim to synchronize timetables in a transit network in order to minimize the passenger waiting time. Omar et al. [9] combine timetabling problems and vehicle scheduling problems in a bi-objective integrated model and use e-constraint method to solve it. Zhang et al. [10] propose a micro-optimization model based on timetables and track maintenance task scheduling. Omar et al. [11] raise the issue of multi-cycle bus schedule optimization to maximize passenger load and minimize vehicle aggregation. They also propose an optimization problem of synchronization of bus schedules [12]. Sun et al. [13] develop three optimization models and designed demand-sensitive schedules by using equivalent time intervals to demonstrate train operation. Shi et al. [14] design an integer linear optimization method to solve the service-oriented train schedule optimization problem of supersaturated subway line collaborative passenger flow control.

Existing research on the timetable of the subway mainly focuses on the optimization of the two goals of subway energy consumption and passenger waiting time. However the exiting studies mainly consider the quality of the train as a fixed parameter. In actual situations, the quality of trains will continue to change as the number of passengers changes. In this paper, the impact of actual passengers on the quality of trains is taken into account, so that the calculation of energy consumption will be more accurate. Moreover, this paper optimizes the timetable through actual passenger demand data, and on this basis, making this two-objective model more in line with real life.

This paper is organized as follows. Section 2 presents the considered problem and formulates its bi-objective optimization model. In Sect. 3, the proposed solution approach is presented. A case study based on Fuzhou Metro Line 1 is conducted to verify the proposed approach. Finally, Sect. 5 draws a conclusion.

2 Problem Description and Formulation

2.1 Problem Description

The research angle of this paper is to consider the change of train quality, so as to make the calculation of energy consumption more accurate. Most of the previous studies set the mass of the train to a constant. In fact, in addition to the mass of the empty train, the mass of the train should also consider the passenger capacity on the train. Based on the actual passenger arrival rate, this study sets the average quality of a person on this basis. By optimizing the arrival and departure times of multiple trains, the actual passenger load data of each train can be obtained. On this basis, the train quality closer to the actual life data can be obtained, thereby making the calculation of energy consumption more accurate.

This study does not consider the second waiting of passengers, because in many second-tier and even third-tier cities, subway development is still in its infancy, and a relatively sound subway line network has not been formed. The number of people who choose subway as a travel tool is usually not enough. It makes the subway full or even overloaded, so it is reasonable for all passengers to get on the train within the time of one waiting without the second waiting.

Furthermore, many studies consider the train departure interval as a constant by default, and optimize on this basis. This study considers that the train departure interval is not fixed but meets certain departure interval constraints. In other words, it can also be understood that the departure interval of every two trains is used as a decision variable. We optimize the arrival and departure time of trains at each station based on the actual passenger arrivals. By changing the relevant specific settings, such as considering changes in quality and changing departure intervals, the total waiting time for passengers, the calculation of the energy consumption of the train is more accurate and meets the actual requirements.

2.2 Formulation

Before formulating the considered problem, the following basic assumptions are made.

1. All passengers wait only once, and there are no passengers waiting twice. This means that in the departure interval of any two trains at the same station, all waiting passengers will board the next train that is about to arrive. This assumption is reasonable in places where there are fewer subway lines, such as the Fuzhou subway system. Because these subway lines have a small passenger flow, they will not be crowded like the Beijing or Shanghai subway system, which will cause two or even three waits.

2. The process of the train traveling between the two stations is composed of a uniform acceleration and two deceleration processes. These are the acceleration phase, taxi phase and braking phase of the train. The accelerations of these three segments are considered constant in this paper.
3. The resistance and additional running resistance of the train during driving are a fixed value.

The notations used for the problem formulation are listed in Table 1.

2.3 Objective Function

(1) Passenger waiting time

According to Hypothesis 1, all passengers wait only once, that is, after arriving at the platform, the passenger will take the next train that will reach the platform. This means that the total waiting time of this passenger will only be determined by the decision variable d_{in} , which is the time the train leaves platform n . Passenger arrival rate $q_n^v(t)$. We will obtain statistics based on the smart card swipe data of the platform. The passenger waiting time function is as follows:

$$W = \sum_{i=2}^I \sum_{n=1}^{N-1} \left[\sum_{t=d_{(i-1)n}}^{d_{in}} \sum_{v=n+1}^N q_n^v(t)(t - d_{(i-1)n}) \right] + \sum_{n=1}^{N-1} \left[\sum_{t=0}^{d_{1n}} \sum_{v=n+1}^N q_n^v(t)(t - d_{1n}) \right] \quad (1)$$

(2) Net energy consumption

The actual mass expression of train i at stations n to $n + 1$ is as follows:

$$m_{in} = P_{in}m_p + m_t \quad (2)$$

The number of passengers between stations is expressed as follows:

If $i = [2, I]$, $n = [2, N - 1]$, then

$$P_{in} = \sum_{k=2}^n \left[\sum_{t=d_{(i-1)k}}^{d_{ik}} \sum_{v=k+1}^N q_k^v(t) - \sum_{v=1}^{k-1} \sum_{t=d_{(i-1)v}}^{d_{iv}} q_v^k(t) \right] + \sum_{t=d_{(i-1)v}}^{d_{iv}} \sum_{v=2}^N q_1^v(t) \quad (3)$$

If $i = 1$, $n = [2, N - 1]$, then

$$P_{in} = P_{1n} = \sum_{k=2}^n \left[\sum_{t=0}^{d_{1k}} \sum_{v=k+1}^N q_k^v(t) - \sum_{v=1}^{k-1} \sum_{t=0}^{d_{1v}} q_v^k(t) \right] + \sum_{t=0}^{d_{11}} \sum_{v=2}^N q_1^v(t) \quad (4)$$

If $i = 1$, $n = 1$, then

Table 1 Parameters used for formulation

Parameter symbol	Meaning of representation
<i>Indices</i>	
I	$i = 1, 2, \dots, N$
N	$n = 1, 2, \dots, N$
<i>Parameters</i>	
m_p	Average quality of people
m_t	The quality of an empty train
$q_n^v(t)$	The passenger who arrived at station n at time t and destination is station v
η_1	Energy conversion efficiency during acceleration
η_2	Energy conversion efficiency during braking
a_1	Acceleration during traction
a_2	Acceleration during taxiing
a_3	Acceleration during braking
r	Basic running resistance
g	Additional force based on lines and curves
$\underline{t}_{(n,n+1)}$	Minimum travel time at station $(n, n + 1)$
$\overline{t}_{(n,n+1)}$	Maximum travel time at station $(n, n + 1)$
\underline{D}	The minimum dwell time on the platform
\overline{D}	The maximum dwell time on the platform
\underline{h}	Minimum departure interval
\overline{h}	Maximum departure interval
l_n	Distance between station n and station $n + 1$
<i>Decision variables</i>	
a_{in}	When train i arrives at station n
d_{in}	When train i leaves station n
<i>Intermediate variables</i>	
m_{in}	The quality of train i between stations $n \sim n + 1$
c_{in}	Switching time point of acceleration and taxi phase of train i between $(n, n + 1)$
b_{in}	Switching time point of train i between $(n, n + 1)$ taxi and braking phases
v_{cin}	Speed of train i between $(n, n + 1)$ acceleration and coasting phase switching time
b_{cin}	Speed of train i at the time point of switching between taxi and braking phases between $(n, n + 1)$
f_{ain}	Traction force of train i running between station n and $n + 1$
f_{bin}	The braking force of train i running between station n and $n + 1$

$$P_{in} = P_{11} = \sum_{t=0}^{d_{11}} \sum_{v=2}^N q_1^v(t) \quad (5)$$

If $i \in [2, I]$, $n = 1$, then

$$P_{in} = P_{il} = \sum_{t=d_{(i-1)1}}^{d_{i1}} \sum_{v=2}^N q_1^v(t) \quad (6)$$

The speed relation expression is as follows:

$$v_{cin} = (c_{in} - d_{in})a_1 \quad (7)$$

$$v_{bin} = (c_{in} - d_{in})a_1 - (b_{in} - c_{in})a_2 \quad (8)$$

$$v_{bin} - (a_{i(n+1)} - b_{in})a_3 = 0 \quad (9)$$

$$v_{cin}^2/(f_a - r + g) + (v_{cin}^2 - v_{bin}^2)/(r - g) + v_{bin}^2/(f_b + r - g) = 2l_n \quad (10)$$

The expressions of traction and braking force for train i traveling between stations n and $n + 1$ are as follows:

$$f_{ain} = m_{in}a_1 + r - g \quad (11)$$

$$f_{bin} = m_{in}a_3 + g - r \quad (12)$$

The expressions of traction energy consumption and regenerative braking energy consumption are as follows:

$$F_{in}(t) = \begin{cases} f_{ain}v_{in}(t)/\eta_1, & t \in [d_{in}, c_{in}) \\ 0, & t \in \text{others} \end{cases} \quad (13)$$

$$J_F = \sum_{i=1}^I \sum_{n=1}^{N-1} \sum_{t=d_{in}}^{b_{in}} F_{in}(t) \quad (14)$$

$$B_{in}(t) = \begin{cases} f_{bin}v_{in}(t)\eta_2, & t \in [b_{in}, a_{i(n+1)}] \\ 0, & t \in \text{others} \end{cases} \quad (15)$$

The calculation expression of overlapping time is as follows:

$$O_1^{f1} = \begin{cases} 0, c_{in} < b_{(i+1)(n-1)} \\ \min\{c_{in} - d_{in}, c_{in} - b_{(i+1)(n-1)}\}, d_{in} \leq b_{(i+1)(n-1)} \leq c_{in} \leq a_{(i+1)n} \\ \min\{c_{in} - d_{in}, a_{(i+1)n} - d_{in}, b_{(i+1)(n-1)} \leq d_{in} \leq a_{(i+1)n} \\ 0, a_{(i+1)n} < d_{in} \end{cases} \quad (16)$$

$$O_2^{f1} = \begin{cases} 0, c_{(i+1)n} < b_{in} \\ \min\{c_{(i+1)n} - d_{(i+1)n}, c_{(i+1)n} - b_{in}\}, b_{in} \leq c_{(i+1)n} \leq a_{i(n+1)} \\ \min\{c_{(i+1)n} - d_{(i+1)n}, a_{i(n+1)} - d_{(i+1)n}, b_{in} \leq d_{(i+1)n} < a_{i(n+1)} \\ 0, a_{i(n+1)} < d_{(i+1)n} \end{cases} \quad (17)$$

The expression of the regenerative braking energy conversion coefficient is as follows:

$$\eta_3 = \frac{\sum_{i=1}^{I-1} \sum_{n=1}^{N-1} (O_1^{f1} + O_2^{f1}) + \sum_{i=1}^{I-1} \sum_{n=N+1}^{2N-1} (O_1^{f2} + O_2^{f2})}{\sum_{i=1}^{I-1} \sum_{n=1}^{N-1} (a_{i(n+1)} - b_{in}) + \sum_{i=1}^{I-1} \sum_{n=N+1}^{2N-1} (a_{i(n+1)} - b_{in})} \quad (18)$$

The expression of net energy consumption is as follows:

$$J = J_F - J_B \eta_3 \quad (19)$$

2.4 Constraints

(1) Travel time constraints between stations

$$\underline{t}_{(n,n+1)} \leq a_{i(n+1)} - d_{in} \leq \overline{t}_{(n,n+1)}, \forall i \in [1, I], n \in [1, N-1] \quad (21)$$

(2) Dwell time constraint

$$\underline{D} \leq d_{in} - a_{in} \leq \overline{D}, \forall i \in [1, I], n \in [1, N-1] \quad (22)$$

(3) Departure interval constraint

$$\underline{h} \leq d_{i1} - a_{(i-1)1} \leq \overline{h}, \forall i \in [2, I], n \in [1, N-1] \quad (23)$$

3 Solution Approach

In this paper, the elitist non-dominated sorting genetic algorithm (NSGA-II) is adapted to efficiently solve the studied problem, and obtain its Pareto frontier for

decision-makers. Here we first give the explanation of Pareto dominance relationship. For a multi-objective problem with n objectives $f_i(x)$, $i = 1, 2 \dots n$, given two feasible solutions. X_a and X_b , if the following two conditions are true, it is said that X_a dominates X_b :

For $\forall i \in 1, 2 \dots n$, there exists $f_i(X_a) \leq f_i(X_b)$;

For $\exists i \in 1, 2 \dots n$, there exists $f_i(X_a) < f_i(X_b)$.

If there is no other decision variable that can dominate a decision variable, then the decision variable is called a non-dominated solution. In a set of solutions, the non-dominated solution Pareto rank is defined as 1, the non-dominated solution is deleted from the solution set, the remaining solution Pareto rank is defined as 2, and so on, and the Pareto rank of all solutions in the solution set can be obtained.

The proposed algorithm is mainly composed of the following steps [15, 16]. First, it randomly initializes the population, then quickly non-dominated sorting, and then calculates the crowding distance. Afterwards, children are generated through the binary tournament, and the parent and child are mixed for crossover and mutation operations, and then all individuals are sorted through fast non-dominated sorting and crowding distance, leaving good individuals until the number of breeding generations is enough to produce several Pareto fronts. The algorithm contains three most important parts, it proposes a fast non-dominated sorting algorithm, uses crowding degree and crowding degree comparison operator and introduces elite strategy. The specific instructions are as follows:

A. Fast non-dominated sorting

Non-dominated sorting algorithm is used to stratify the population of size n , the specific steps are as follows:

- (1) Set $i = 1$;
- (2) For all $j = 1, 2, \dots, n$, and $j \neq i$, according to the above definition, compare the dominant and non-dominant relationship between individual x_i and individual x_j ;
- (3) If there is no individual x_j better than x_i , then x_i is marked as a non-dominated individual;
- (4) Let $i = i + 1$ and go to step (2) until all non-dominated individuals are found;

The set of non-dominated individuals obtained through the above steps is the first-level non-dominated layer of the population. Then, the marked non-dominated individuals are not considered (that is, these individuals will not be compared for the next round), and then steps (1)–(4) are re-run to get the second-level non-dominated layer. And so on, until the entire population is stratified. The fast non-dominated sorting of NSGA-II reduces the computational complexity.

B. Crowding degree computing

The crowding degree estimation is to obtain the crowding degree estimation of the solution around the specific solution in the population. We calculate the average

distance between the two points on both sides of this point according to each objective function. This value is used as an estimate of the perimeter of the cuboid with the nearest neighbor as the vertex (called the crowding factor). The calculation of the crowding coefficient needs to sort the population according to the ascending order of the size of each objective function. Therefore, for each objective function, the boundary solution is specified as the value of the infinite distance. All other intermediate solutions are specified as equal to the normalized absolute difference between two adjacent solutions. The calculation method is the same for other objective functions. All the congestion coefficient values are calculated by summing the distance values of each target of the individual, and each objective function will be normalized before calculating the congestion coefficient. Specific steps are as follows:

The congestion degree i_d at each point is set to 0;

For each target, the population is sorted non-dominated, so that the crowding degree of the two individuals at the boundary is infinite, that is $i_d = \sum_{j=1}^m (|f_j^{i+1} - f_j^{i-1}|)$.

The formula to calculate the crowding degree of other individuals is as follow.

$$i_d = \sum_{j=1}^m (|f_j^{i+1} - f_j^{i-1}|) \quad (24)$$

After the previous fast non-dominated sorting and crowding degree calculation, each individual i in the population has two attributes: the non-dominated sorting order i_{rank} and the crowding degree determined by the non-dominated sort i_d . As long as any of the following conditions are true, the individual i wins.

$$i_{rank} < j_{rank} \quad (25)$$

$$i_{rank} = j_{rank}, \quad \text{and} \quad i_d > j_d \quad (26)$$

C. Elite strategy

First, the new population Q_t produced by the t_{th} generation is merged with the parent P_t to form R_t , and the population size is $2N$. Then R_t performs non-dominated sorting, generates a series of non-dominated sets Z_i and calculates the degree of congestion. Since the offspring and parent individuals are included in R_t , the individuals contained in the non-dominated set Z_1 after non-dominated sorting are the best in R_t , so first Z_1 is put into the new parent population P_{t+1} . If the size is less than N , continue to fill the next level of non-dominated set Z_2 into P_{t+1} until the size of the population exceeds N when Z_3 is added. For individuals in Z_3 use the crowding comparison operator to make the number of P_{t+1} individuals reach N . Then generate new progeny populations Q_{t+1} through genetic operators (selection, crossover, mutation).

4 Computational Results

A. Basic data

In this study, we use data from Fuzhou Metro Line 1. Fuzhou Metro Line 1 has 21 platforms. The distance between each platform has been actually measured. Passenger arrival rate data is provided by Fuzhou Metro. This paper takes the data of the morning peak as an example to study the optimization problem of ten trains, on the basis of which the effectiveness of the model and algorithm is verified.

The original timetable is presented in Table 2. The departure interval of each car in the morning and peak period is fixed, and the departure interval is 405 s.

The proposed NSGA-II is coded in C++ under Visual Studio 2019. All the computational experiments are conducted on the personal computer with an Intel Core i5 2.2 GHz CPU and 4 GB RAM. The values of input parameters are presented in Table 3.

B. Optimal timetable

According to the above NSGA-II algorithm and model, several Pareto optimization solutions are obtained. The initial timetable is compared to verify the effectiveness of the model and algorithm. Table 3 lists the comparison between several Pareto solutions and the initial timetable.

As show in Table 4, passenger waiting time of the current timetable is 669,863 s, and the energy consumption is 7,535,778,493 Nm. The average energy consumption

Table 2 Current timetable of Fuzhou metro line 1

Station	1	2	3	4	5	6	7
Arrival time (s)	0	150	300	450	600	750	900
Departure time (s)	30	180	330	480	630	780	930
Station	8	9	10	11	12	13	14
Arrival time (s)	1050	1200	1350	1500	1650	1800	1950
Departure time (s)	1080	1230	1380	1530	1680	1830	1980
Station	15	16	17	18	19	20	21
Arrival time (s)	2100	2250	2400	2550	2700	2850	3000
Departure time (s)	2130	2280	2430	2580	2730	2880	-

Table 3 The input parameters

Parameter	N	I	m_t	m_p	a_1	a_2	a_3	r	g
Value	21	10	311,800	60	1	0.02	0.8	2000	500
Unit	-	-	kg	kg	m/s ²	m/s ²	m/s ²	N	N
Parameter	η_1	η_2	pop_size	max_gen					
Value	0.7	0.3	50	100					

Table 4 Comparison between the initial and obtained timetables

	Solution number	Passenger waiting time	Energy consumption
Optimal timetable	1	704,977	690,636,039
	2	685,769	692,496,655
	3	682,912	693,777,691
	4	652,443	696,101,564
	5	639,362	699,650,248
	6	634,991	703,475,147
	7	468,649	878,454,146
	8	725,823	528,732,081
Average		641,421	697,915,446
Current timetable		669,863	753,578,493
Reduction rate		4.2%	7.4%

of the optimized Pareto solution is 697,915,446 Nm, which reduces 7.4% compare with the current timetable. The average Passenger waiting time of the optimized Pareto solution is 669,863 s, which reduces 4.2% compare with the current timetable. With one of the best found Pareto solution, like solution number 5, both objective functions are optimized to a certain degree. The passenger waiting time reduces 4.6%, and energy consumption reduces 7.2%.

These results indicate that the optimal timetable is more reliable and effective. This also proves the effectiveness of the models and algorithms presented above.

5 Conclusion

The main contribution of this paper is to consider the impact of specific passenger flow on train quality as a consideration, so that the quality of the train is regarded as a variable value that changes with the number of passengers. In this way, the calculation of the energy consumption of the train will be more accurate. On this basis, this paper developed a dual-objective model, taking the total passenger waiting time and the energy consumption of trains as the optimization goals, and finally realized the optimization of the timetable. After presenting the model and algorithm, using the example of Fuzhou Metro Line 1, through the assistance of the program, the optimization of the example was finally achieved. The experiment finally obtained a set of Pareto solutions for decision-makers to choose. In this way, passenger waiting time was reduced by 4.2% and train energy consumption was reduced by 7.4%.

In this paper, we consider the optimization of a subway line in a single direction. Of course, this type of research can be extended to bidirectional or even comprehensive optimization of the subway network in future research [16]. In future research, other objective functions can also be added, such as consideration of environmental impact,

maintenance of trains, Robustness, etc. In some hypothetical settings, it can be more reasonable. This paper considers that there are no passengers waiting twice, but in some large cities, due to the huge number of passengers, it is very likely that the second or even three waits will appear. All of the above will be considered in more depth in future research.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China under Grants 71701049, 71871059 in part by the Humanities and Social Science Foundation of the Chinese Ministry of Education under Grant 21YJA630096, in part by the 2nd Fujian Young Eagle Program Youth Top Talent Program, and in part by Fujian science and technology economic integration service platform.

References

1. Yang X, Ning B, Li X, Tang T (2018) A two-objective timetable optimization model in subway systems. *IEEE Trans Intell Transp Syst* 15(5):1913–1921
2. Yan F, Goverde RMP (2019) Combined line planning and train timetabling for strongly heterogeneous railway lines with direct connections. *Transp Res Part B Methodol* 127:20–46
3. Yang X, Wu J, Sun H, Gao Z, Yin H, Qu Y (2019) A bi-objective timetable optimization model for urban rail transit based on the time-dependent passenger volume. *IEEE Trans Intell Transp Syst* 20(2):604–615
4. Yang X, Chen A, Wu J, Gao Z, Tang T (2018) An energy-efficient rescheduling approach under delay perturbations for metro systems. *Transportmetrica B* 7(2):1–15
5. Yang X, Chen A, Ning B, Tang T (2016) A stochastic model for the integrated optimization on metro timetable and speed profile with uncertain train mass. *Transp Res Part B Methodol* 91:424–445
6. Yang S, Lia F, Wu J, Timmermans HJP, Sun H, Gao Z (2020) A bi-objective timetable optimization model incorporating energy allocation and passenger assignment in an energy-regenerative metro system. *Transp Res Part B* 133:85–133
7. Zhou W, You X, Fan W (2020) A mixed integer linear programming method for simultaneous multi-periodic train timetabling and routing on a high-speed rail network. *Sustainability* 12(3):11–31
8. Abdolmaleki M, Masoud N, Yin Y (2020) Transit timetable synchronization for transfer time minimization. *Transp Res Part B* 131:143–159
9. Ibarra-Rojas OJ, Giesen R, Rios-Solis YA (2014) An integrated approach for timetabling and vehicle scheduling problems to analyze the trade-off between level of service and operating costs of transit networks. *Transp Res Part B* 70:35–46
10. Zhang Y, D'Ariano A, He B, Peng Q (2019) Microscopic optimization model and algorithm for integrating train timetabling and track maintenance task scheduling. *Transp Res Part B* 127:237–278
11. Ibarra-Rojas OJ, López-Irarragorri F, Rios-Solis YA (2016) Multiperiod synchronization bus timetabling. *Transp Sci* 50(3):805–822
12. Ibarra-Rojas OJ, Rios-Solis YA (2012) Synchronization of bus timetabling. *Transp Res Part B Methodol* 46B(5):599–614
13. Sun L, Jin JG, Lee D-H, Axhausen KW, Erath A (2014) Demand-driven timetable design for metro services. *Transp Res Part C* 46:284–299
14. Shi J, Yang L, Yang J, Gao Z (2018) Service-oriented train timetabling with collaborative passenger flow control on an oversaturated metro line: an integer linear optimization approach. *Transp Res Part B* 110:26–59

15. Deb P, Meyarivan A (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* 6(2):182–197
16. Yang X, Wu J, Sun H, Gao Z, Yin H, Qu Y (2019) Performance improvement of energy consumption, passenger time and robustness in metro systems: a multi-objective timetable optimization approach. *Comput Indust Eng* 137:106076

Wind Tunnel Tests on Aerodynamic Noise from the Head Car of a High-speed Train



Hai Deng, Yigang Wang, Qiliang Li, Zhe Shen, Yang Gao, and Jin Lisheng

Abstract Characteristics of aerodynamic noise is always difficult to extract by both experimental and numerical methods. Based on the aero-acoustic wind tunnel test results of a 1:8 scale high-speed trains model with three cars, the properties of the aerodynamic noise sources of the head car were analyzed. In view of the fact that dipole source is the main aerodynamic noise source of the high-speed train, the acoustic similarity relationship between the different models was deduced. Subsequently, the aerodynamic noise properties of the 1:1 scale model (approximate full-scale train) were analyzed. The results show that the main noise sources of the 1:1 scale model are the bogie regions and the noise energy of 80 Hz–4 kHz is dominant at the train speed of 250 km/h. The noise is the strongest at the frequency of 160–400 Hz, and the noise of the nose tip and the obstacles is the same order as that of the bogie area after 2.5 kHz. The acoustic similarities between the different models are related to the third power of the speed, and the first power of the model length, the propagation distance and the frequency. The noise difference between the different models is related to scale size and frequency. The frequency at full scale has an exponential relationship with the scale model's frequency, not following a liner relationship with the scale.

H. Deng

School of Traffic and Transportation Engineering, Central South University, Changsha 410083, China

Y. Wang (✉) · Q. Li · Z. Shen

Shanghai Automotive Wind Tunnel Center, Tongji University, Shanghai 201804, China

e-mail: yigang.wang@sawtc.com

Shanghai Key Lab of Vehicle Aerodynamics and Vehicle Thermal Management Systems, Shanghai 201804, China

H. Deng · Y. Gao

General Research and Development Department, CRRC Changchun Railway Vehicles Co. Ltd., Changchun 130062, China

J. Lisheng

School of Vehicle and Energy, Yanshan University, Changchun 130022, China

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,

Lecture Notes in Electrical Engineering 775,

https://doi.org/10.1007/978-981-16-5429-9_12

Keywords High-speed train model · Aero-acoustic wind tunnel · Aerodynamic noise · Acoustic similarity

1 Introduction

The high-speed train head shape in the design, on one hand, highlights the modelling features and reflects the unique brand culture of train manufacturers; on the other hand, it is the first component of the train encountering the air, combined with such complex curve structures, i.e., the nose tip, the cowcatcher, the bogie area, which makes the aerodynamic and noise problem even more important and attracts more attentions in academic communities. One of key issues is the research on aerodynamic noise of the head car in high-speed train design. This is because: (1) The aerodynamic noise energy of high-speed trains, characterized by dipole, increases with the speed to the sixth power [1] and higher than that of other cars, due to the higher velocity around the head car. (2) the high-speed airflow passes through the nose tip, the cowcatcher and the bogie area, and the complexity of the bogie region contributes it to the one of main noise sources of high-speed trains [2–4]. The noise has effects on the quiet environment along the railway line, and in the carriages, particularly in the driver cabin. How to design a low noise high-speed train becomes of importance.

Although numerical simulation and wind tunnel test for aerodynamic noise of high-speed trains have been being carried out constantly [5–8], limited by the large scale and unsteady calculation, massive computational resource is required in simulations; and numerical methods is not fully developed still so that the accuracy and the competency of the result are not ensured; therefore, it is hardly to simulate the aerodynamic noise characteristics at full scale. The wind tunnel tests on aerodynamics of high-speed trains have been carried out abroad [4, 9–11], but strictly on specified projects, because of expensive model-making and test cost. The aero-acoustic wind tunnel test is even rare. In wind tunnel tests of high-speed trains, the 1/8 scaled model is always used. Besides, smaller scale models, e.g., 1/15 and 1/20, are sometimes adopted. The study on exploring aero-acoustic similarity rules to solve the issue about how to transfer the results at scaled model to the full scale has been carrying out, but lack of effective aero-acoustic similarity rules makes it hard to be achieved.

In this paper, wind tunnel tests were carried out, by means of the sound pressure and sound source identification measurement technology. A 1/8 scaled three-car grouping high-speed train model was used to analyse the distributions of aerodynamic noise sources and the frequency characteristics. According to the established similarity rules in different scaled models, the aerodynamic noise characteristics of a full-scale train model were discussed. This study can provide a reference for understanding aerodynamic noise from the head car of a high-speed train, and is also an exploration on the aero-acoustic similarity problems.

2 Experiment

2.1 Test Measuring System and Analysis Method

The tests were carried out in the aero-acoustic wind tunnel of the Shanghai Automotive Wind Tunnel Center in Tongji University. The dimensions of the test section are $27\text{ m} \times 17\text{ m} \times 12\text{ m}$ with a jet port of 27 m^2 . The acoustic performance index completely meets the test requirements as compared to those similar wind tunnels in the world. When the flow speed at the nozzle is 250 km/h , the background noise of the test section is 73.4 dB(A) , and the cut-off frequency is 50 Hz . The test section is shown in Fig. 1.

In order to retain the structural features of the head, middle and tail cars of a high-speed train and meet the test section of the tunnel, a $1/8$ scaled three-car grouping model was used as the test model which was fixed on an aerodynamic floor (including a corresponding track). The shape of the head and tail cars is the same, as shown in Fig. 2.

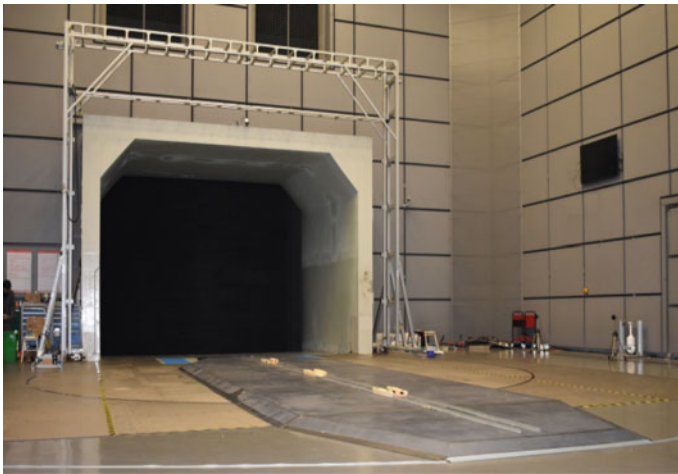
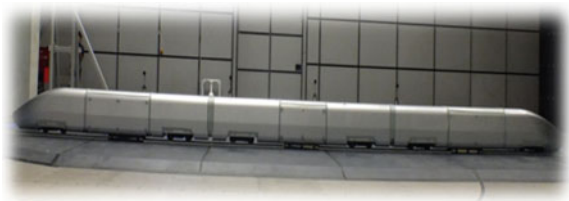


Fig. 1 Wind tunnel platform with floor for model installation on ground

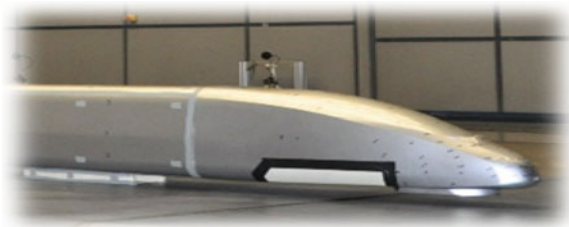
Fig. 2 $1/8$ scaled high-speed train model with three cars



For this high-speed train model, the main noise sources are distributed in the nose tip of the head car (including cowcatcher), bogies, connecting part between cars and airflow separation position of the tail car, etc. Current measuring methods, such as far-field microphone and sound source identification techniques, are difficult to completely distinguish the mentioned sound sources. Therefore, in this study, two models were adopted in the tests. One is to smooth the bottom of the model, i.e., removing the bogie and filling the cavity with covers to retain the same cross-section as that of train body. The gaps are all covered with tape, as shown in Fig. 3. In addition, four supporters were used to connect the train model and the floor, and each two were enclosed by an airfoil body to guide the flow, Fig. 4. The inter-carriage gaps were also covered with curved sheets with tape to smooth the train body. All those done aims to study the aerodynamic noise at the nose tip and cowcatcher (named the semi-closed model). The second is to use a relatively real train model with bogies and cavities to identify the noise sources of the head car (named the real model).

The location of measuring points is shown in Fig. 5. Four free-field microphones were distributed along the length of the model at 7.5 m away from the central axis of the train model, with an interval of 3 m. Those microphones that lie above the ground at 0.8 m were used to measure the noise radiated from the train body. The microphone array was arranged 4.2 m away from the central axis of the train model and close to the head car, which is parallel to the centerline of the model and is used to orient different noise sources of the head car.

Fig. 3 Smooth treatment on train body



(a)



(b)

Fig. 4 Supporters wrapped with airfoil guide: **a** installation, **b** cylinder and airfoil

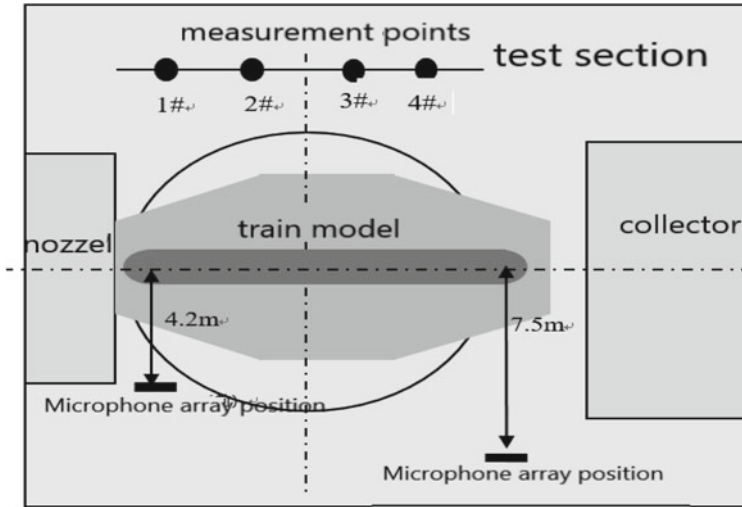


Fig. 5 Distribution of measuring points

2.2 Results and Analysis

2.2.1 Far-Field Noise Characteristics of the Three-Car Grouping Model

Figure 6 shows the noise spectrum measured at points 1#–4# at the wind speed of 250 km/h. The measurement point 1# is close to the head car, and the point 4# (near

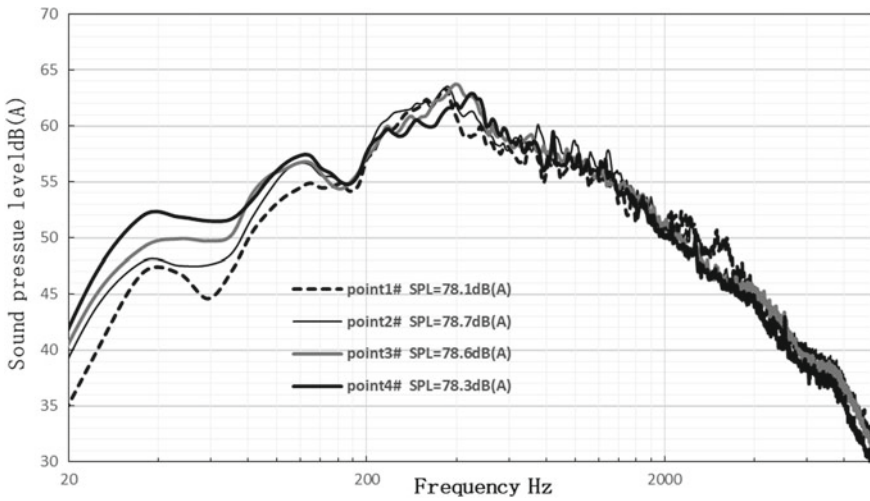


Fig. 6 Noise spectrum comparison for measuring points 1#–4#

the tail car) is far from the head car. The observation is that when the frequency is less than 200 Hz, the low frequency noise increases gradually with the point moving towards the tail car and wind tunnel collecting port, which indicating that the head car does not contribute to low frequency noise. From the research results of noise source identification, the noise mainly comes from the wind tunnel collecting mouth and tail car separation region. Because the low-frequency compression wave generated by the jet flow at the nozzle and the diversion structure at the wind tunnel collecting port propagating upstream, is the largest at the wind tunnel collecting port, and the low frequency shedding vortex separates from the tail car, both emerge to form strong low frequency acoustic characteristics in this region. According to the following analysis, the aerodynamic noise of the bogie shows a dominant role in 100 Hz–5 kHz. In Fig. 6, the noise at point 1# is also mainly from the bogie, and shows a small difference as compared to the other points. In the range of 700 Hz–2 kHz, point 1# has a smaller value, indicating that besides of the two bogies of the head car, the bogies of the middle and tail cars have different contributions to the three measuring points. However, in 2–4 kHz, the sound pressure increase at point 1#; this is because the contributions mainly come from the nose tip and cowcatcher where the separation is produced and leads to flow fluctuations. Those fluctuations are the main source of noise. In addition, the low frequency characteristics probably exist in this region, and they may be covered by the bogie regional noise. Based on the above analysis, the point 1# presents a different trend as compared to others. This difference is induced by the head car, since the point is close to it. Therefore, the noise law at point 1# can be used for the analysis of the noise characteristics of the head car.

2.2.2 Noise Characteristics of the Head Car Region

According to above analysis, the noise at point 1# can represent the noise radiated characteristics of the head car. Therefore, the results of different models at point 1# were analysed in Fig. 7, in order to approximately understand the aerodynamic noise characteristics of each part of the head car. There are four cases, i.e., the empty wind tunnel (EWT), the empty wind tunnel with the installation of track (EWTT), the semi-closed model (SCM) and the real model (RM). The maximum wind speed is 250 km/h in the wind tunnel tests.

When the floor is installed in the empty wind tunnel, the sound pressure at point 1# is 71.9 dB(A) that is lower than 73.4 dB(A) in the empty wind tunnel (there is additional aerodynamic noise from the moving ground system in the tunnel), indicating that the floor brings no additional noise into the system. After the track is installed, the additional noise only increases in the high frequency range of 1–10 kHz, showing the typical noise characteristics from the track under the impact of the flow, which can be changed when the train body is fixed on ground. While the semi-closed model is put in the tunnel, the noise increases in 50 Hz–1 kHz, and this noise probably comes from the supporters according to later analysis. Small peaks come out in 1–5 kHz. Those can be attributed to the flow separation around the nose

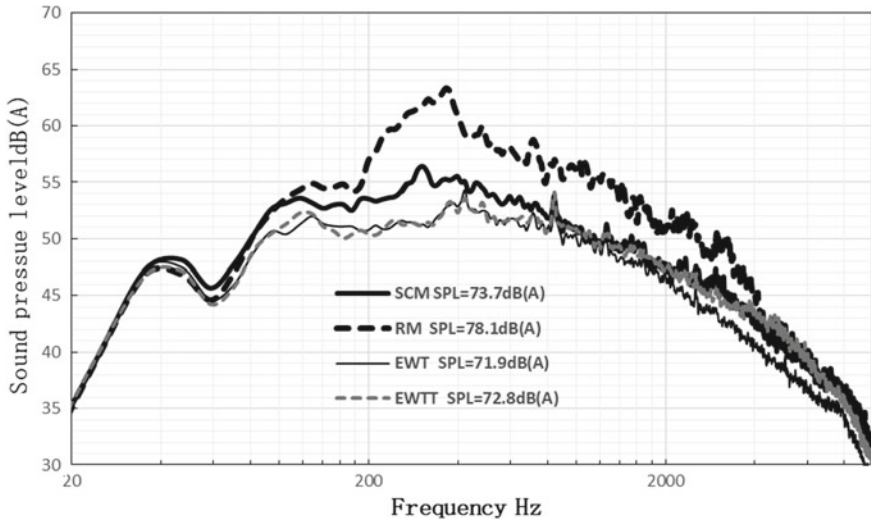


Fig. 7 Noise spectrum comparison for wind tunnel with floor, track, semi closed model and real vehicle model

tip and cowcatcher. The corresponding total sound pressure is 73.7 dB(A) that is a little higher than 72.8 dB(A) in the EWTT case. When the real model is fixed on the floor, the noise energy increases significantly in 100 Hz–5 kHz, particularly in 200–500 Hz. The total sound pressure reaches 78.1 dB(A). Thus, the bogie region becomes the main noise source of the head car. The noise sources in this region are complex, which include the near-ground cavity noise (Bogies in the cavities), air impact noise, etc., showing broad frequency characteristics.

2.2.3 Noise Source Identification of the Head Region of Car

In the last section, the noise characteristics of the measuring point 1# is the sum of all noise sources from a three-car grouping model, can only approximately shows the noise radiation characteristics of the head car. To understand the noise characteristics of the head car, the microphone arrays were used to identify each sound source location and its frequency characteristic.

(1) Noise characteristics of the semi-closed model

The semi-closed model retains the nose tip, the cowcatcher and the curved surface, but the bogie region has been smoothed. Between the two bogies there are airfoil supporters. Therefore, the main noise sources come from those regions. As shown in Fig. 8, the microphone array can identify the location of a sound source and its corresponding frequency range. The blue line in Fig. 8a is the linear spectrum of the sound source (contours) in the Fig. 8b, and the step figure is 1/3 octave spectrum. The

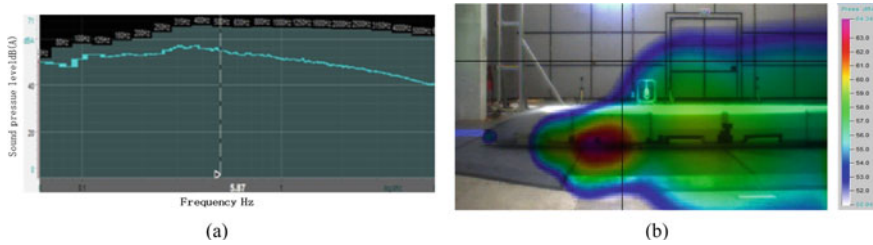


Fig. 8 Sound sources identification at full frequencies for head car model (left: Noise spectrum; right: noise contour)

black region represents the frequency range corresponding to the sound source in Fig. 8b which represents the location of the sound source. The color ranges from red to blue, representing the sound pressure levels from large to small. Figure 8 shows the main noise source characteristics of the semi-closed model in the full frequency band. The largest noise source is located at the supporting part of the airfoil (red zone), followed by the second loudest source, which is mainly wind tunnel background noise, and the third largest noise source is located at the nose tip and cowcatcher.

In addition to the airfoil supporter and wind tunnel background noise, the nose tip and cowcatcher become the first noise sources of the model.

Figure 9 shows the distribution of noise sources in different frequency bands. In the low frequency range, the background noise of the wind tunnel shows a dominant role, but it is not one of the main noise characteristics from the model; therefore, no more discussion is conducted about it. According to the analysis of the noise source contours, it is observed that in 1/3 octave the noise from the airfoil supporters is dominant in the center frequency 250–800 Hz; that from the nose tip starts at 1 kHz, and continues up to 5 kHz. Since the nose tip and cowcatcher are adjacent each other and the resolution of the sound source identification array is not sufficient, so it is not easy to distinguish noise from the nose tip and cowcatcher. Therefore, it

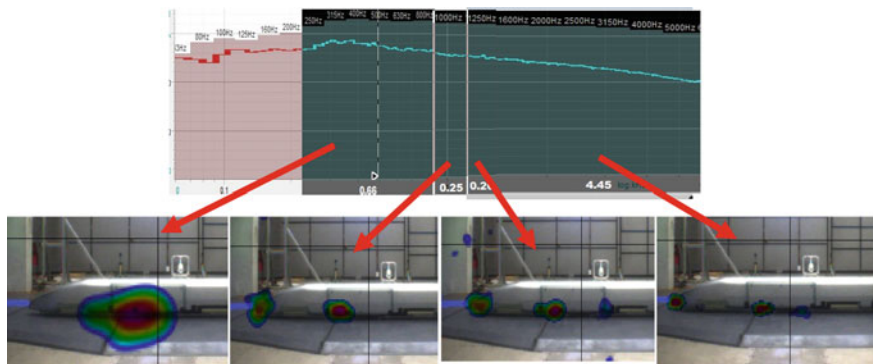


Fig. 9 Sound sources identification at different frequency bands for head car model

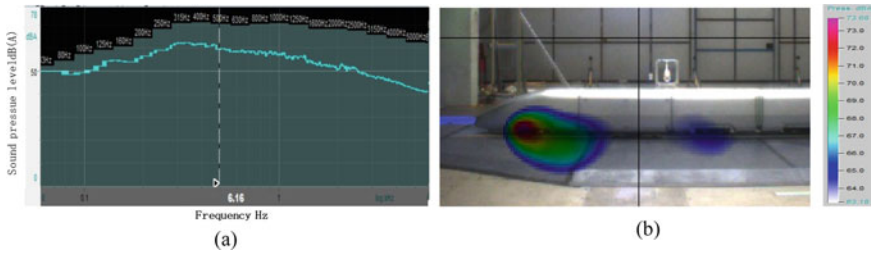


Fig. 10 Sound sources identification at full frequencies for head car model (left: noise spectrum; right: noise contour)

can be concluded that the nose tip and cowcatcher generate high frequency noise. Whether the low frequency characteristic exists or not is hard to be verified, because this characteristic is perhaps covered by the noise from the supporters. The noise produced by the train body and streamlined head is much lower than that of the nose tip and cowcatcher.

(2) Noise characteristics of the real model

The semi-closed model highlights the noise characteristics from the nose tip and cowcatcher, and further confirm that the low noise from the train body. In this section, a real model with bogies was used to understand the noise characteristics.

Figure 10 shows the characteristics of the main noise sources from the real model in all frequencies. It can be seen that the largest sound source located at the first bogie region of the head car (red area), and the second largest one is from the second bogie region (blue area). The difference between both is 8 dB(A). It can be attributed to the bottom flow, i.e., the velocity in the first bogie region is faster than that in the second one. As we know, the aerodynamic noise, characterized by dipole, increases with the speed to the sixth power. Meanwhile, for the cavity noise, the increase of flow speed leads to more energy, and therefore stronger noise. The noise from the nose tip and cowcatcher is not shown in the contour, indicating it is very low and does not have the same order as that from the bogie.

The noise distribution in different frequencies is shown in Fig. 11. In the low frequency range, the background noise of the wind tunnel still shows a dominant role, but no more discussion is conducted about it. According to the analysis of the noise source contours in Fig. 11, it is observed that in 1/3 octave the noise comes from the front of the second bogie is dominant in the center frequency of 200–250 Hz. Although the velocity is lower than that in the first bogie region, but the noise increases; therefore, why this happens needs to be clarified. However, based on the analysis of Fig. 6, when the frequency is higher than 100 Hz, the noise from the bogies plays a key role, which shows a difference. Due to the poor resolution of the sound source identification equipment in low frequency, the noise source is difficult to distinguish when the frequency is less than 200 Hz; so the noise frequency of the second bogie should be lower, at least to 100 Hz. After that, the noise from the first bogie shows a dominant role in 250–500 Hz, and until to 800 Hz. As the frequency

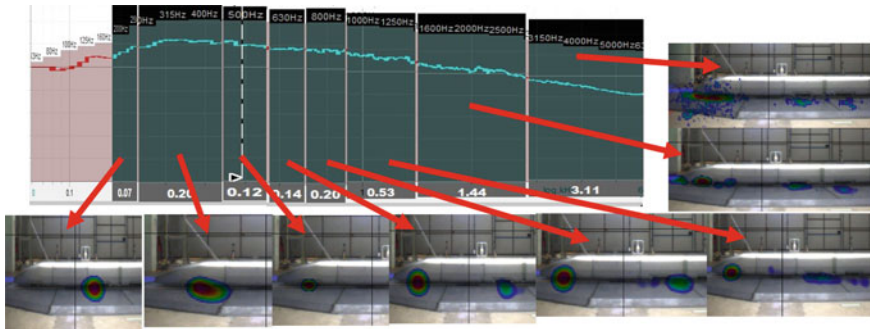


Fig. 11 Sound sources identification at different frequency bands for head car model

increases, the position of the sound source moves forward. The higher velocity in the first bogie region contributes to the noise energy, while the reason for the source moving forward probably comes from the increasing high frequency noise produced by the impact of air on the small components of the front wheels of the first bogie. In 630 Hz, the second noise source is shown in the second bogie region, which can be concluded that this source indeed exists from the frequency of 200 Hz and covered by the stronger noise from the first bogie. Starting from 1 kHz, the third noise source comes out in the third bogie region (i.e., the first bogie of the middle car). With the frequency increasing from 1 kHz, the noise sources in the nose tip, the cowcatcher, and the second and third bogies show stronger as compared to the rest. The high frequency noise comes from the air interaction with train components from the nose tip to the cowcatcher and till each bogie. Also, the air impact on small components of wheels is one of the main reasons.

The spectral characteristics of the noise sources identified by the microphone array are basically consistent with those in Sects. 1.2.1 and 1.2.2 using far-field measuring technology, but more details can be concluded as: (1) the strongest noise source of the head car is the first bogie region, and the frequency distribution (Center frequency in 1/3 octave) covers from 250 Hz to 5 kHz. The rear part of the bogie region presents the low frequency noise characteristics, while the high frequency characteristics show in the front. (2) The second bogie is the second noise source with the frequency distribution from 100 Hz to 5 kHz, but in this region the distribution of noise is slightly complex. (3) In 100 Hz–5 kHz, the noise energy in the bogie region shows a dominant role, and the strongest is in 200–Hz. (4) The nose tip and the cowcatcher is the third noise source with far less energy but high frequency characteristics as compared to the bogie regions. When the frequency increases to 3.15 kHz, it can be the same order as that in the bogie.

3 Derivation of Noise Characteristics of Different Scale Models

The noise characteristics of a 1/8 scale high-speed train model cannot fully represent those of a real train. Based on the dipole source for the properties of main noise characteristic of high-speed trains, the aero-acoustic similarity related to different scale models was established, and the spectrum characteristics of a full-scale model were derived from the current data of a 1/8 scale model, which, to a certain extent, can show the noise characteristics of the real train.

According to Ref. [12], the sound power radiated outward from the dipole sound source is defined as:

$$W_D = \frac{4\pi r^2 k^2 b^2}{3\rho_0 c_0} P_{eD}^2 \quad (1)$$

where, W_D is the sound power radiated outward from the dipole sound source. r is the distance from the source to the observation point. b is the distance between two pulsating spheres of the dipole. $k = \omega/c_0$ is the number of the wave. $\omega = 2\pi f$ is the angular frequency and f is the frequency. ρ_0 and c_0 are the density of the air and sound velocity, respectively. P_{eD} is the sound pressure on the envelope surface. The aerodynamic noise produced in the process of motion is related with the speed, the characteristic dimension and the media properties. Thus, the sound power of the dipole sound source can be changed as [13]:

$$W_D = k_d \frac{\rho_0 U^6 L^2}{c_0^3} \quad (2)$$

where, k_d is the coefficient. U is the flow velocity. L is the characteristic dimension. P_{eD} is obtained from Eq. (1), and it can be changed in terms of the sound pressure level.

$$SPL = 10 \log_{10} \frac{P_{eD}^2}{P_{ref}^2} = 10 \log_{10} M_D \frac{\rho_0 U^6 L^2}{r^2 f^2} \quad (3)$$

where, $P_{ref} = 2 \times 10^{-5}$ is the reference pressure.

$$M_D = \frac{3k_D \rho_0^2}{64 \times 10^{-10} \pi^3 b^2}$$

According to Eq. (3), the difference of the noise pressure level between the 1/8 scale and full scale three-car grouping model can be calculated as:

$$\Delta SPL = SPL_1 - SPL_8$$

Table 1 Sound pressure level comparison between measurement and calculation at 1#

Velocity (km/h)	Test (dB(A))	Calculation (dB(A))	Error (dB(A))
175	67.9		
200	71.6	71.4	0.2
225	75	74.4	0.6
250	78.1	77.2	0.9

$$\begin{aligned}
 &= 10 \log_{10} \frac{M_D}{M_D} + 20 \log_{10} \frac{U_1^3 L_1 r_8 f_8}{U_8^3 L_{81} r_1 f_1} \\
 &= 20 \log_{10} \frac{U_1^3 L_1 r_8 f_8}{U_8^3 L_{81} r_1 f_1} \quad (4)
 \end{aligned}$$

where, the subscript 1 stands for the full-scale model and the subscript 8 is for the 1/8 scale model. This equation shows the conversion relationship between SPL of the two scale models. To validate the correct of Eq. (4), a comparison between the test and calculation of the sound pressure level at point 1# was conducted, as shown in Table 1. Using the data at the velocity of 175 km/h as a benchmark, the sound pressure level at other velocities can be calculated. As shown in the table, the error is less than 1 dB(A), indicating that the precision of Eq. (4) is sufficient.

To convert the noise characteristics of the 1/8 scale model to those of the full scale model, let $U_1 = U_8$, $r_1 = r_8$, $L_1 = 8L_8$ in Eq. (4), so

$$\Delta SPL = SPL_1 - SPL_8 = 18 + 20 \log_{10} \frac{f_8}{f_1} \quad (5)$$

According to the results of the wind tunnel test, the sound pressure level at point 1# is 78.1 dB(A). Because of $f_1 < f_8$, the sound pressure level at point 1# for a full-scale model is $SPL_1 > 96.16$ dB(A). Thus, the frequency relationship between two models can be shown as:

$$f_1 = f_8 10^{\frac{18 - \Delta SPL}{20}} \quad (6)$$

To obtain the relationship between the frequencies of models, the sound pressure level from the real train should be listed. According to the Ref. [14], at the train speed of 250 km/h, the sound pressure level at 25 m away from the centerline of the train is between 84 and 90 dB(A), and the averaged value is 87 dB(A). Supposed the aerodynamic noise takes up the main component, the averaged can be approximately considered as the aerodynamic noise. According to the acoustic attenuation rule, the noise at the corresponding 7.5 m is 98 dB(A). Thus, $\Delta SPL_1 > 19.9$ dB(A). Then the relationship can be changed as

$$f_1 = 0.8 f_8 \quad (7)$$

Combined with the noise frequency characteristics of 1/8 scale model in wind tunnel tests and Eq. (7), the noise characteristics of the full-scale model can be deduced: (1) The strongest noise source of the head car is the first bogie region, and the frequency distribution (Center frequency in 1/3 octave) covers from 200 Hz to the whole frequency band. (2) The second bogie is the second noise source with the frequency distribution from 80 Hz to the whole frequency band. (3) In 80 Hz–4 kHz, the noise energy in the bogie regions shows a dominant role, and the strongest is in 160–400 Hz. (4) The nose tip and the cowcatcher is the third noise source with high frequency characteristics as compared to the bogie regions. When the frequency increases to 2.5 kHz, it can be the same order as that in the bogie.

4 Conclusions

Wind tunnel tests on the aerodynamic noise of 1/8 scale three-car grouping high-speed train models have been carried out. The experimental data at far-field regions, of different structure models and of noise identification were used to study the main noise sources, distributions and frequency characteristics produced from the head car. Given the features of the dipole source from the aerodynamic noise source of the high-speed train, the aero-acoustic similarity related to different scale models was derived, and the noise characteristics of the head car at full scale was analysed. The conclusions can be drawn as:

- (1) The strongest noise source of the head car is the first bogie region, and the noise energy covers all frequency bands. The rear part of the bogie region presents the low-medium frequency noise characteristics, while the high-medium frequency characteristics show in the front. The second bogie is the second noise source, and the frequency distribution is generally identical to the first bogie but presents slightly complex. For the real model, in 80 Hz–4 kHz, the noise energy in the bogie regions shows a dominant role, and the strongest is in 160–400 Hz. The nose tip and the cowcatcher are the third noise source with far less energy but high frequency characteristics as compared to the bogie regions. When the frequency increases to 2.5 kHz, it can be the same order as that in the bogie.
- (2) Aero-acoustic similarity relationship between different models with the third power of the speed of the model, with the liner of the length of the model, propagation distance and frequency of the wave. The increase of the noise is mainly related to the dimensions of the model and the corresponding frequency. The noise frequency of the full-scale model is lower than that of the 1/8 scale model, further indicating the relationship with the dimensions of the model, but the main noise sources and the distribution characteristics of both models are similar. It has confirmed that the relationship of noise frequencies in different scale models, as shown in Eq. (6), is not a simple rule of multiple, but an exponential relationship.

References

1. Zuo C, Ming L, Peng H (2007) Detection methods and counter measures for aerodynamic sound sources generating from rolling stock. *Foreign Rolling Stock* 44(6):35–40
2. Kitagawa T, Nagakura I (2000) Aerodynamic noise generated by Shinkansen cars. *J Sound Vib* 231(3):913–924
3. Tai Otte C, Gautier PE, Thompson DJ et al (2003) Identification, modeling and reduction potential of railway noise sources: a critical survey. *J Sound Vib* 267(2):447–468
4. Nagakura K (2006) Localization of aerodynamic noise sources of Shinkansen trains. *J Sound Vib* 293:547–556
5. Thompson DJ, Latorre IE, Liu XW et al (2015) Recent developments in the prediction and control of aerodynamic noise from high-speed trains. *Int J Rail Transp* 3(3):119
6. Zhu CL, Hemida H, Flynn D et al (2017) Numerical simulation of the slipstream and aeroacoustic field around a high-speed train. *Proc Inst Mech Eng Part F J Rail Rapid Transit* 231(6):740
7. Zhu J, Wang Y, Yang Z et al (2017) Effect of bogie fairing on the flow and aerodynamic noise behaviour around the high-speed train bogie. *J Tongji Univ (Nat Sci)* 45(10):1512
8. Zhu J, Ren L, Lei Z (2018) Effect of bogie cavity on flow and flow-induced noise behaviour around high-speed train bogie region. *J Tongji Univ (Nat Sci)* 46(11):1556
9. Yamazaki N, Takaishi T, Toyooka M et al (2008) Wind tunnel tests on the control of aeroacoustic noise from high speed train. *Noise Vibr Mitigation Rail Transp Syst* 2008:33–39
10. Lauterbach A, Ehrenfried K, Kröber S et al (2010) Microphone array measurements on high-speed trains in wind tunnels. In: Berlin beamforming conference
11. Nagakura K (2001) Method of analyzing the wind tunnel test data measured with directional microphone systems. *Q Rep RTRI* 42(2):104–109
12. Farassat F (2000) Introduction to generalized functions with applications in aerodynamics and aeroacoustics. *J Sound Vib* 230:460–462
13. Ma D (2004) *Modern acoustic theory foundation*. Science Press
14. Zhang S (2009) Noise mechanism, source identification and control for a 350km/h high-speed train. *China Railway Sci* 30(1):86–91

Invulnerability Simulation of Urban Agglomeration Traffic Network Cascading Failure Under Random Attack Strategy



Chengbing Li, Zhicheng Yang, and Zhenyu Liu

Abstract Aiming at the cascading failure problem in urban agglomeration traffic network, this paper proposed a new cascading failure model. In this model, we divided the node states into normal, pause, and failure with the judgment criterion of the degree of node overload. For the pause node we adopted the flow distribution rule of residual capacity priority, and for the failure node, the flow distribution was based on the node distance and degree of connectivity. Meanwhile, this paper also considered the enduring ability between nodes or edges in the load transfer process. We studied the effects of overload capacity parameter which means the node redundancy capacity, passenger flow adjustment which means the initial network load, and the impact of different attack strategies on urban agglomeration traffic network invulnerability with the simulation of actual passenger flow in Hu-Bao-E urban agglomeration. The results show that the overload capacity parameter has a threshold of 1.2 with the node attacks, so the invulnerability would not be improved when we increase the value of overload capacity parameter. The influence of passenger flow adjustment parameter on network invulnerability shows different characteristics under different attack patterns. Compared with the attack edge, the attack node has a more significant impact on traffic network invulnerability, because node bears large passenger flow, and it will easily be cascading failure when it is attacked. And in the attack process, it shows the cliff-like decline phenomenon of correction relative scale of maximum connected subgraph, and the urban agglomeration traffic network has an extremely strong scale-free characteristic.

Keywords Urban agglomeration · Comprehensive transportation · Cascading failure · Simulation · Invulnerability

C. Li (✉) · Z. Liu

Transportation Institute, Inner Mongolia University, Hohhot 010021, China
e-mail: bingbingnihao2008@126.com

Z. Yang

School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

C. Li

Inner Mongolia Engineering Research Center for Urban Transportation Data Science and Applications, Hohhot 010021, China

1 Introduction

The urban agglomeration has gradually replaced the signal city as the basic region unit of global competition and division in the process of urbanization [1]. In the urban agglomeration, the traffic network composed of various traffic modes is becoming increasingly complex, and the network maintenance is facing more and more challenges. Once mishandled, the invulnerability of urban agglomeration traffic network will greatly reduce in the face of unexpected faults or disasters.

In terms of complex networks, recent studies have gradually shifted from the characteristics of independent networks to interdependent networks [2–4]. Buldyrev et al. gave a framework for the analysis of cascading effects of interdependent networks with comparing and analyzing the different effects of cascading failure on independent networks and interdependent networks [5]. In addition, scholars also studied the improvement of invulnerability and the restoration strategy optimization of the interdependent networks in cascading failure [6, 7]. Presently, cascading failure is mainly studied from two aspects of network dynamics and topology. Danziger et al. discussed the cascading dynamics mechanism of interdependent networks [8], and Yin et al. studied the influence of node capacity parameters on scale-free network invulnerability in cascading failure and found that the threshold of node capacity was negatively correlated with the network invulnerability [9]. From different perspectives, scholars have established different cascading failure models, including probabilistic cascade failure model based on the initial load and capacity parameter distribution [10], and probabilistic heterogeneous capacity allocation model based on capacity redistribution mode, which simulation results show that capacity allocation mode has a great impact on network robustness [11]. Li et al. proposed the capacity allocation model based on the node degree value and found it has a high robustness [12]. Ren et al. proposed the load distribution of cascade failure model based on the residual capacity of nodes, and it shows that this model can effectively reduce the scale of cascade failure [13]. Meanwhile, some scholars have studied the influence of different attack modes on network robustness [14–16].

Network cascading failure has been studied extensively in reality [17–19]. In terms of traffic, the invulnerability of urban public transportation network is getting more attention [20–22], especially for the rail transit network [23]. However, there is a lack of research on the mechanism of cascading failure in urban agglomeration traffic network. And most researches study the propagation of cascading failure in traffic network from the perspective of topology, which is difficult to describe the actual network state and lacks practical significance. In addition, it is necessary to be noticed that the urban agglomeration traffic network system is a complex network system coupled by various transportation modes, the dependence between different traffic subnets is very strong, and the propagation of cascade effect exists widely among traffic subnets and in the internal subnets. Therefore, we should pay much attention on the construction of comprehensive traffic network when build cascading failure model. Compared with previous studies, the differences of this paper are mainly reflected in the following points.

- (1) Based on the traditional capacity-load model, we define three node states as normal, pause and failure, and give the state judgment criterion based on the node overload probability.
- (2) Due to cascading effect, we consider the edge enduring ability between nodes and node enduring ability between edges in the case of load transfer.
- (3) We give different traffic allocation strategies for nodes in different states by considering actual traffic network. And we propose a new concept of network connectivity distance weight effect, which describes the tendency of failure nodes allocated flow to peripheral nodes.
- (4) From the perspective of urban agglomeration, we use actual passenger flow to simulate the impact of initial load and capacity adjustment parameters on the traffic network invulnerability.

The content of this paper is as follows. Firstly, we establish a weighted traffic network model of urban agglomeration based on the actual traffic flow and our previous researches [24, 25]. Secondly, we establish cascading invulnerability model of urban agglomeration traffic network and correction load-capacity model, and give the network attack strategies and measurement index of invulnerability. Thirdly, we take an example of Hu-Bao-E urban agglomeration for simulation. Lastly we give the conclusions of study.

2 The Weighted Traffic Network Model of Urban Agglomeration

2.1 The Construction of Urban Agglomeration Traffic Network Model

The construction of existing traffic network models is usually based on the road mapping or site mapping method, and this paper builds traffic network model of urban agglomeration based on the site mapping method considering that the regional scope of urban agglomeration is wide and the road distribution density is large. All stations of different traffic modes in the urban agglomeration are regarded as nodes of the network, all lines connecting the stations are regarded as edges of the network, and the stations or edges which geographical location is close are combined to form composite nodes or edges. Taking daily passenger volume of station as the node weight and daily passenger volume of edge as the edge weight, we construct the weighted traffic network model of urban agglomeration. Based on our previous researches [24, 25], we make the following assumptions and definitions.

Assumption 1: We do not consider the direction of traffic network. If a node can reach another node, it is assumed that another node can also return to the node, so the traffic flow is approximately same in two directions. That is, the urban agglomeration traffic network is a non-directional network.

Assumption 2: The traffic supply of each line in the urban agglomeration traffic network is determined by the actual traffic demand. And traffic demand is stable in a short time.

Definition 1: Traffic subnet. The traffic network of road, railway, water transportation and aviation in urban agglomeration is defined as the traffic subnet.

Definition 2: Compound node. We superimpose two or more traffic subnets and combine the bus station, railway stations, ports and airports which geographical location is close with ignoring geographical distance to get the compound node. If passengers' transfer walking time between stations is in the acceptable and reasonable range, we think the geographical location of the station is close. V' is the set of all combined nodes, and V'' is the set of compound nodes.

Definition 3: Compound edge. After the traffic subnets are superimposed, if there are multiple edges connected between two nodes, it will be regarded as one edge connected, and this edge is called composite edge. E' is the set of all combined edges, and E'' is the set of compound edges.

Therefore, we construct the traffic network model of urban agglomeration by connecting multiple traffic subnets of different types with compound nodes and compound edges, denoted by $G(V, E, W, H)$. Among them, V is the nodes set of the urban agglomeration traffic network, $V = \cup_{s=1}^q V_s \setminus V' \cup V'' = \{v_1, v_2, \dots, v_i, \dots, v_n\}$. q is the number of compound traffic subnets, and n is the number of nodes in compound network. E is the edges set corresponding to V , $E = \cup_{s=1}^q E_s \setminus E' \cup E'' = (e_{ij})_{n \times n}$. When $V_i R V_j$, $e_{ij} = 1$, and when $V_i \bar{R} V_j$, $e_{ij} = 0$. The weight of compound edge is the sum of combined edges weights of every traffic subnets, so $W = (w_{ij})_{n \times n}$, where w_{ij} is the weight of edge ij . And the weight of compound node is the sum of combined nodes weights of every traffic subnets, so $H = \{h_1, h_2, \dots, h_i, \dots, h_n\}$, where h_i is the weight of node i .

2.2 Correction of Traffic Network Weight

Because of the uniqueness of traffic demand, the traffic demand of urban agglomeration shows certain characteristics in time. The passenger transport volume is different in different lines and different time segment, and in order to describe this feature of urban agglomeration traffic demand, we introduce the parameter of $\&$.

$w_s^f = \& \cdot r_s^f$ is the edge weight of edge f of traffic subnet s in urban agglomeration. Among them, $\&$ is the number between $[0, 1]$, which is used to describe the difference of passenger volume in different periods or adjust the passenger volume of network. And r_s^f is the number of daily providing seats of edge f of traffic subnet s . We can get the node weight corresponding to edge weight, and the formulas are as followings.

$$w_s^f = \& \cdot r_s^f \quad (1)$$

$$h_s^i = \sum_{f=1}^{k_s^i} \cdot r_s^f \quad (2)$$

where k_s^i is the number of connected edges of node i .

3 The Construction of Cascading Invulnerability Model of Urban Agglomeration Traffic Network

3.1 Correction of Load-Capacity Model

The load-capacity model of urban agglomeration traffic network is as followings.

$$c_i = r * l_i$$

$$c_{ij} = r * l_{ij}$$

where c_i is the capacity of node i , c_{ij} is the capacity of edge ij , l_i is the initial load of node i , l_{ij} is the initial load edge ij , and r is the overload capacity parameter of urban agglomeration traffic network.

Assuming that the initial load of the node v_i is the node weight h_i before it is attacked, and the initial load of edge ij is also the edge weight w_{ij} . Therefore, $l_i = h_i$, $l_{ij} = w_{ij}$.

3.2 Recognition of Node and Connected Edge State

(1) The recognition of node state is as formula (3).

$$if \begin{cases} l_i < c_i & normal \\ l_i \geq c_i \text{ and } rand > p_i & pause \\ l_i \geq c_i \text{ and } rand \leq p_i & failure \end{cases} \quad (3)$$

where $rand$ is the random number between 0 and 1, and p_i is the failure probability of node i , which definition formula (4) is as following.

$$p_i = \frac{l_i - c_i}{c_i} \quad (4)$$

(2) The recognition of connected edge state is as formula (5).

$$if \begin{cases} l_{ij} < c_{ij} & normal \\ l_{ij} \geq c_{ij} \text{ and } rand > p_{ij} & pause \\ l_{ij} \geq c_{ij} \text{ and } rand \leq p_{ij} & failure \end{cases} \quad (5)$$

where $rand$ is the random number between 0 and 1, and p_{ij} is the failure probability of edge ij , which definition formula (6) is as following.

$$p_{ij} = \frac{l_{ij} - c_{ij}}{c_{ij}} \quad (6)$$

3.3 The Load Redistribution of Overload Node and Connected Edge

(1) The load redistribution of pause node

For the pause node, the demand exceeds supply of traffic facilities because the line formed by overload node and peripheral node is still open to traffic as usual. We consider that the passengers will preferentially choose the line with larger residual capacity of the connected edges in this condition. Therefore, we can expand the cascading effects by further distributing the superfluous load of overload node to adjacent nodes through edges. So we construct the load redistribution model based on the residual capacity of the edge, as formulas (7) and (8).

$$\Delta e_{ij(t+1)} = \Delta l_i \frac{C_{ij} - l_{ij}}{\sum_{a \in \Gamma_i} (C_{ia} - l_{ia})} \quad (7)$$

$$\Delta l_i = l_i - c_i \quad (8)$$

where $\Delta e_{ij(t+1)}$ is the load distributed from pause node i to edge ij at time $t + 1$, and Γ_i is the set of neighbor nodes of node i .

(2) The load redistribution of failure node

For the failure node, all the lines connected to it will become invalid, which will lead to the traffic paralysis. When passengers choose the route, the first factor to be considered is the accessibility to the destination station, and the second important factor for transfer passengers is the station connectivity in the whole network. So we construct the index of connected distance weight effect by considering distance and connectivity in the whole network. $f_{(ij)}$ is used to describe the connected distance weight effect, as formulas (9) and (10).

$$f_{(ij)} = \frac{B_{ij}}{\sum_{f \in \Gamma_i} B_{if}} / \frac{d_{ij}}{\sum_{f \in \Gamma_i} d_{if}} \quad (9)$$

$$B_{ij} = \frac{\sum_{l \neq p \in V} \sigma_{lp}^{(e_{ij})}}{\sum_{l \neq p \in V} \sigma_{lp}} \quad (10)$$

where B_{ij} is the betweenness of edge ij , $\sigma_{lp}^{(e_{ij})}$ is the number of shortest paths between node v_l and v_p through the edge e_{ij} , σ_{lp} is the number of shortest paths between node v_l and v_p , Γ_i is the set of neighbor nodes before node i is failure, and d_{ij} is the actual space distance of edge ij . So we construct the load redistribution model of failure node, as formulas (11) and (12).

$$\Delta e_{ij(t+1)} = \Delta l_i \cdot f_{(ij)} \quad (11)$$

$$\Delta l_i = l_i \quad (12)$$

where $\Delta e_{ij(t+1)}$ is the load distributed from failure node i to node j at time $t + 1$.

(3) The load redistribution of overload connected edge

For the pause edge, we distribute the overload to two adjacent nodes equally, and assume that Δv_i is the load which pause edge distributes to node v_i , as formula (13).

$$\Delta v_i = \frac{c_{ij} - l_{ij}}{2} \quad (13)$$

For the failure edge, we distribute the entire load to two adjacent nodes, and assume that Δv_i is the load which failure edge distributes to node v_i , as formula (14).

$$\Delta v_i = \frac{l_{ij}}{2} \quad (14)$$

3.4 The Random Attack Strategy

Traditional attack strategies on nodes or edges in network can be divided into proactive and random strategies. For the proactive attack, we firstly sort the nodes or edges from high to low according to its importance, then successively attack them and observe the load distribution in network. For the random attack, we number the nodes or edges in the whole network and attack them randomly. The previous researches just attack the nodes or edges unilaterally. However, in urban agglomeration traffic network, the failure of nodes and edges has great uncertainty because the network coverage area is large and running distance is long. Therefore, this paper attacks the nodes or edges of network randomly. Then, we distribute the load of attacked node to adjacent edges according to vested distribution strategy, and delete

the node and its connected edges. And then we distribute the load of attacked edge to connected nodes and delete the attacked edge.

3.5 The Measurement Index of Invulnerability

The existing researches of invulnerability mostly use the maximum connected subgraph to measure the network invulnerability, which means to compare the ratio of the number of effective nodes in the maximum connected subgraph before and after the attack. However, because of the large network coverage area, long running distance and the superposition of various transportation networks in urban agglomeration, there are a large number of nodes and edges in urban agglomeration with the complex situation and large differences of passenger volume between nodes. So the conventional measurement index will no longer be applicable to urban agglomeration transportation network. Therefore, this paper corrects the measurement index, and the new index is the ratio of the sum of all effective nodes weight in the maximum connected subgraph before and after the attack. We use the S to describe the relative scale of maximum connected subgraph, as formula (15).

$$S = \frac{\sum_{i=1}^{N'} h_i}{\sum_{i=1}^N h_i} \quad (15)$$

where N' is the number of nodes in maximum connected subgraph after the attack, and N in the number of nodes without attack.

3.6 The Cascading Invulnerability Model of Urban Agglomeration Traffic Network

Based on the above analysis, we construct the cascading invulnerability model of urban agglomeration traffic network, and the steps are showing as follows.

Step 1: The initialization of network. We construct the compound traffic network $G(V, E, W, H)$. And determine the passenger flow adjustment parameter α , overload capacity parameter r , the initial load and capacity of common connected edges and nodes in network, and the initial load and capacity of compound connected edges and nodes in network.

Step 2: Attacking the nodes or edges randomly from global perspective. When we attack the node, it is failure and we delete the failure node and its connected edges. And when we attack the edge, it is failure and we delete the failure edge.

Step 3: The distribution of load. For attacked node, we determine the proportion of the load distributed to peripheral nodes with formulas (9) and (10), and determine

the load distributed to adjacent nodes at the next time with formulas (11) and (12). For attacked edge, we distribute the load to adjacent nodes equally.

Step 4: Updating the network node load and determining the overload nodes. We judge the node state with formulas (3) and (4), and calculate the set of failure nodes V_u^t at time t , the set of pause nodes V_p^t , and the set of normal nodes V_n^t .

Step 5: The judgment of cascading failure occurrence. If $V_u^t \cup V_p^t = \emptyset$, there is no cascading failure and go to the step 9. If $V_u^t \cup V_p^t \neq \emptyset$, there is occurrence of cascading failure and go to the step 6.

Step 6: The redistribution of load. For pause node, we determine the load distributed to the adjacent edges with formulas (7) and (8). For failure node, we determined the proportion of the load distributed to peripheral nodes with formulas (9) and (10), and determine the load distributed to adjacent nodes at the next time with formulas (11) and (12). For pause edge, we determine the load distributed to the adjacent nodes with formula (13). For failure edge, we determine the load distributed to the adjacent nodes with formula (14).

Step 7: Updating the network load and determining the overload nodes and edges. We judge the node state with formulas (3) and (4), and calculate the set of failure nodes V_u^{t+1} at time $t + 1$, the set of pause nodes V_p^{t+1} , and the set of normal nodes V_n^{t+1} . We judge the edge state with formulas (5) and (6), and calculate the set of failure nodes E_u^{t+1} at time $t + 1$, the set of pause nodes E_p^{t+1} , and the set of normal nodes E_n^{t+1} .

Step 8: The judgment of existing cascading failure. If $V_u^{t+1} \cup V_p^{t+1} \cup E_u^{t+1} \cup E_p^{t+1} = \emptyset$, there is no cascading failure and go to step 9. If $V_u^{t+1} \cup V_p^{t+1} \cup E_u^{t+1} \cup E_p^{t+1} \neq \emptyset$, there is cascading failure and go to step 6.

Step 9: The output S which is the correction relative scale of maximum connected subgraph in urban agglomeration compound traffic network.

Step 10: The judgment of attack end. If $V_n^{t+1} \cup E_n^{t+1} = \emptyset$, go to step 11, otherwise, go to step 2.

Step 11: The end of attack. We get the curve of the correction relative scale of maximum connected subgraph.

4 Example Simulation

4.1 The Construction of Traffic Network Model in Hu-Bao-E Urban Agglomeration

Based on the cascading invulnerability model of urban agglomeration traffic network, this paper simulates as the example of Hu-Bao-E urban agglomeration. Because the Hu-Bao-E urban agglomeration is located in the inland areas of northwest China without water transportation, has a single aviation transportation network, and bears a small proportion of passenger and cargo transportation in the urban agglomeration, we superposed the road and rail traffic subnets to construct the compound traffic

network model of the urban agglomeration. We constructed the road traffic subnet G_1 and rail traffic subnet G_2 with the nodes of all bus and railway stations and the edges of lines in urban agglomeration. Among them, $|V_1| = 156, |V_2| = 29, |E_1| = 429, |E_2| = 90$, and further construct the compound traffic network model of the urban agglomeration. Among them, $|V'| = 11, V'' = 5, V = V_1 \cup V_2 \cup V'' \setminus V', |V| = 179, |E'| = 6, E'' = 3, E = E_1 \cup E \cup E'' \setminus E', |E| = 516$. The compound nodes are shown as Table 1, and the compound traffic network model is shown as Fig. 1.

Table 1 Compound nodes

Normal nodes	The type of traffic network	Compound nodes	The walking time of passenger transfer
Hohhot coach station Passenger western station	Road traffic network	Hohhot coach station Passenger western station-railway station	5 min
Hohhot railway station	Railway traffic network		
Ordos bus station	Road traffic network	Ordos bus station-railway station	8 min
Ordos railway station	Railway traffic network		
Dalad Banner bus station	Road traffic network	Dalad Banner bus station-railway station	6 min
Dalad Banner railway station	Railway traffic network		
Chasuqi bus station	Road traffic network	Chasuqi bus station-railway station	3 min
Chasuqi railway station	Railway traffic network		
Baotou coach station	Road traffic network	Baotou coach station-east railway station	8 min
Baotou east railway station	Railway traffic network		

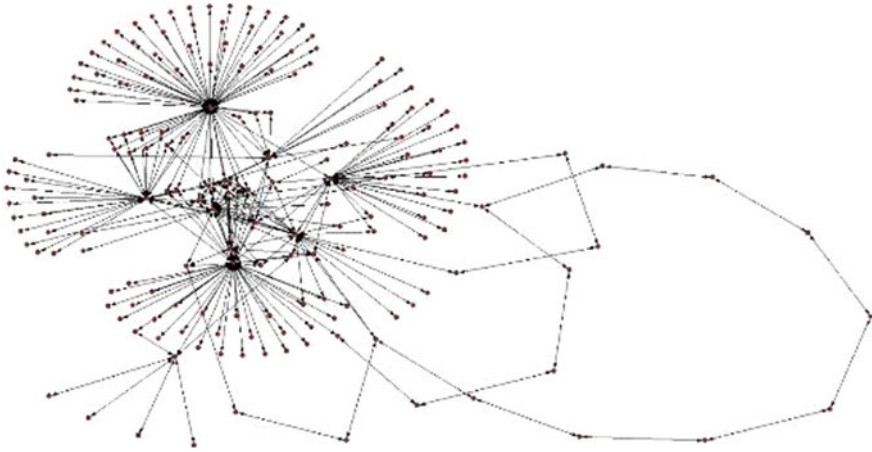


Fig. 1 Hu-Bao-E urban agglomeration road-rail traffic network model

4.2 The Analysis on the Impact of the Overload Capacity Parameter on Urban Agglomeration Traffic Network Invulnerability Under Cascading Failure

4.2.1 The Simulation of Node Invulnerability Under Random Attacks

We need to keep other parameters are unchanged when we study the impact of the overload capacity parameter on urban agglomeration traffic network invulnerability. And the impact of the overload capacity parameter on the urban agglomeration traffic network invulnerability is reflected by the difference of urban agglomeration traffic network invulnerability caused by the change of overload capacity parameter. Based on the assumption of equal supply and demand of urban agglomeration traffic network, this paper assumes that the lines and stations of urban agglomeration traffic network are under the full load operation every day, indicating that the actual passenger volume is equal to the number of actually provided seats. So we consider the impact of the overload capacity parameter on urban agglomeration traffic network invulnerability with the passenger flow adjustment parameter $\alpha = 1$.

We increased the overload capacity parameter r of the compound traffic network with keeping $\alpha = 1$. And we attacked the nodes randomly until the network is global paralysis and results are shown in Fig. 2 and Table 2.

It can be seen from Fig. 2 and Table 2 that the correction relative scale of maximum connected subgraph was 1 when the Hu-Bao-E urban agglomeration compound traffic network was not attacked, and the correction relative scale of maximum connected subgraph was gradually decreasing to 0 with the increasing number of attack nodes, which the network was paralysis. During the experiment, the overload capacity parameter value was uniformly increased from 1.1 to 1.6, but the network invulnerability did not increase uniformly and fluctuated in a certain range. When the

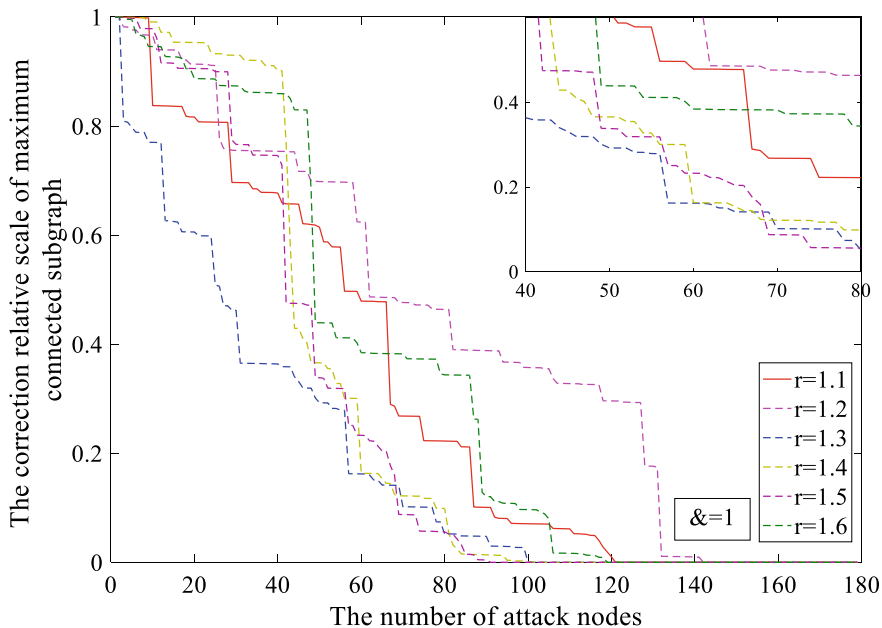


Fig. 2 Overload capacity parameter contrasts effect on invulnerability under random attack

Table 2 Experimental results and analysis

Overload capacity parameter	Accumulative number of nodes in global network paralysis	The average correction relative scale of maximum connected subgraph
1.1	121	0.4507
1.2	144	0.5308
1.3	102	0.3181
1.4	103	0.4771
1.5	91	0.5018
1.6	118	0.5049

overload capacity parameter $r = 1.2$, the network invulnerability was the optimal, the number of accumulative nodes in global network paralysis was 144, and the average correction relative scale of maximum connected subgraph was 0.5308. Compared with $r = 1.1$, the cumulative number of attack nodes increased by 23, the average correction relative scale of maximum connected subgraph increased by 0.0801, and the network invulnerability was significantly improved. When we continued to increase the overload capacity parameter value, the network invulnerability was decreased significantly, the number of accumulative nodes decreased to 102, and the average correction relative scale of maximum connected subgraph decreased

to 0.3181. After that when we continued to increase the overload capacity parameter value, the average correction relative scale of maximum connected subgraph gradually increased, and the network invulnerability was increasing gradually.

In the experiment, we noticed that when the whole network was global paralysis, the cumulative number of attack nodes was 91 with $r = 1.5$. And the value of average correction relative scale of maximum connected subgraph was 0.5018 with $r = 1.5$, indicating that the network had strong invulnerability during the attack. In addition, the cliff-like decline phenomenon of correction relative scale of maximum connected subgraph appeared successively in the simulation process. These all show that the urban agglomeration traffic network has a strong scale-free characteristic and a few key stations play a crucial role in maintaining the network invulnerability, which have a huge impact on urban agglomeration traffic network once they are destroyed.

4.2.2 The Simulation of Edge Invulnerability Under Random Attacks

Based on the consistent traffic lines supply and demand, we kept the passenger flow adjustment parameter $\alpha = 1$, and increased the value of overload capacity parameter r to consider its impact on the urban agglomeration traffic network invulnerability. The simulation results are shown as Fig. 3.

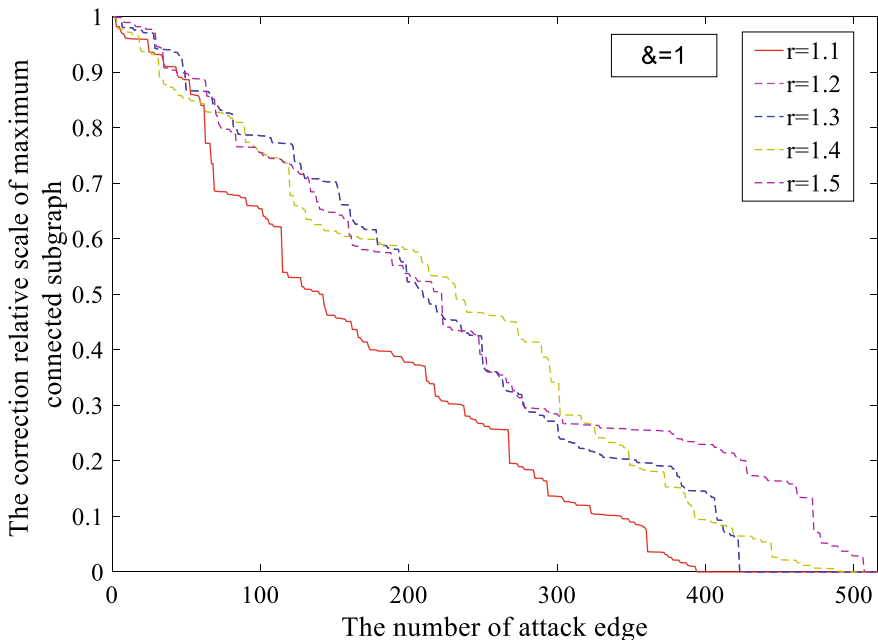


Fig. 3 Overload capacity parameter contrasts effect on invulnerability under random attacks

As can be seen from Fig. 3, when overload capacity parameter $r = 1.1$, the network had the worst invulnerability, and only attacked 396 edges resulting in global network paralysis. When overload capacity parameter $r = 1.2$, $r = 1.3$, $r = 1.4$, $r = 1.5$, the differences of the correction relative scale of maximum connected subgraph were small, and the network invulnerability was almost same. In the middle stage of attack, when $r = 1.4$ in the process of attacking the 250th to 300th edges, the network had strong invulnerability. At the end of the attack, when $r = 1.5$, the network showed stronger invulnerability than the other three cases. In this case, when the cumulative number of attack edges reached 516, the network would be paralysis. The number of attack edges was equal to the number of edges existing in the urban agglomeration traffic network, indicating that there was no cascading failure in the attack process, and the network invulnerability was extremely strong because of large value of overload capacity parameter.

4.3 The Analysis on the Impact of Passenger Flow Adjustment Parameter on Urban Agglomeration Traffic Network Invulnerability Under Cascading Failure

4.3.1 The Simulation of Node Invulnerability Under Random Attacks

When we were studying the impact of passenger flow adjustment parameter on urban agglomeration traffic network invulnerability, the overload capacity parameter was selected as 1.2. The impact of the passenger flow adjustment parameter on the urban agglomeration traffic network invulnerability is reflected by the difference of urban agglomeration traffic network invulnerability caused by the change of passenger flow adjustment parameter. And we attacked the nodes randomly until the network is global paralysis and results are shown in Fig. 4 and Table 3.

It can be seen from Fig. 4 and Table 3 that the change of accumulative number of nodes to make the global network paralysis and average correction relative scale of maximum connected subgraph is extremely uneven with the uniform decrease of passenger flow adjustment parameter from 0.9 to 0.4. When $\alpha = 0.9$, the accumulative number of nodes was largest as 144, and the average correction relative scale of maximum connected subgraph was 0.5271. When $\alpha = 0.8$, the accumulative number of nodes decreased to 91, and the average correction relative scale of maximum connected subgraph increased to 0.5483, which had increased by 0.0212. When $\alpha = 0.7$, the accumulative number of nodes was 102, and the average correction relative scale of maximum connected subgraph was 0.3181 which was minimum. When $\alpha = 0.4, 0.5$, and 0.6, the accumulative number of nodes was decreasing gradually, and the average correction relative scale of maximum connected subgraph was fluctuated between 0.4771 and 0.5120.

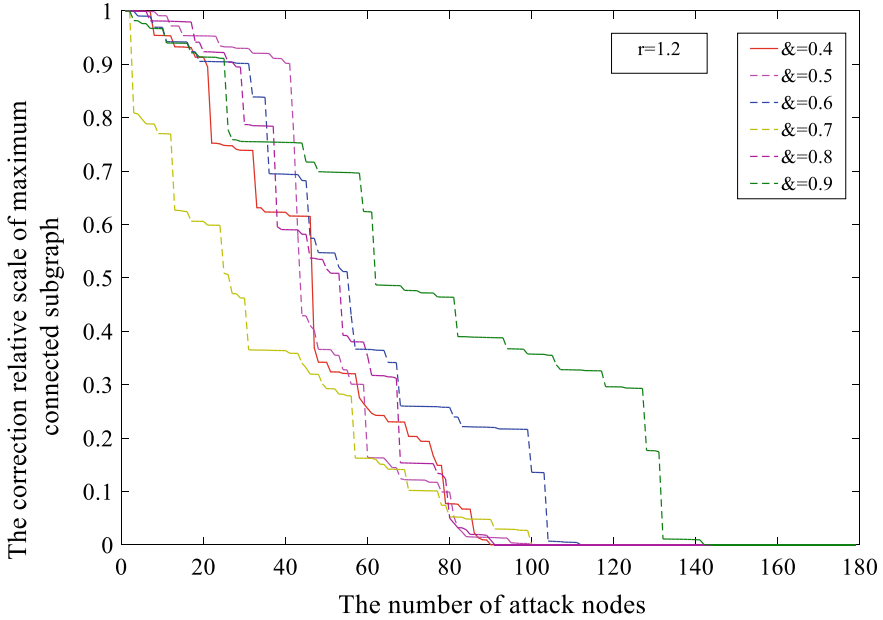


Fig. 4 Passenger flow adjustment parameter contrasts effect on invulnerability under random attacks

Table 3 Experimental results and analysis

Passenger flow adjustment parameter &	Accumulative number of nodes in global network paralysis	The average correction relative scale of maximum connected subgraph
0.4	91	0.5026
0.5	103	0.4771
0.6	112	0.5120
0.7	102	0.3181
0.8	91	0.5483
0.9	144	0.5271

4.3.2 The Simulation of Edge Invulnerability Under Random Attacks

When the overload capacity parameter $r = 1.1, 1.2,$ and $1.3,$ we considered the impact of different passenger flow adjustment parameter values on urban agglomeration traffic network invulnerability. The experiment found that the impact of different passenger flow adjustment parameter values on urban agglomeration traffic network invulnerability was significantly different with $r = 1.1.$ But when $r = 1.2$ and above, the experimental results were not clear. So for the convenient analysis, we selected the overload capacity parameter $r = 1.1$ to investigate the impact of different

passenger flow adjustment parameter values on urban agglomeration traffic network invulnerability, as shown in Fig. 5.

It can be seen from Fig. 5 that network invulnerability was the worst with $\alpha = 0.9$, and because only 430 connected edges attacked can lead to the global network paralysis, the network was more sensitive to attacks, and the correction relative scale of maximum connected subgraph decreased quickly. In the early and late stages of the attack, the network invulnerability was similar with $\alpha = 0.7$ and $\alpha = 0.8$, and the difference of correction relative scale of maximum connected subgraph was small. In the middle stage, the network invulnerability was strong with $\alpha = 0.8$ in the process of attacking the 150th to 350th edges. In addition, the cumulative number of attack edges reached 516 when the network was global paralysis with $\alpha = 0.7$. At this time, the network cascading failure was weak, indicating that the failure of single line did not have a great impact on the whole network operation. And it has a weak impact on the network, because in the urban agglomeration traffic network, the connected edge supports small nodes, bears small passenger flow, and after the cascading failure, this passenger flow of line can be absorbed by the adjacent nodes, which is difficult to produce large-scale cascading failure.

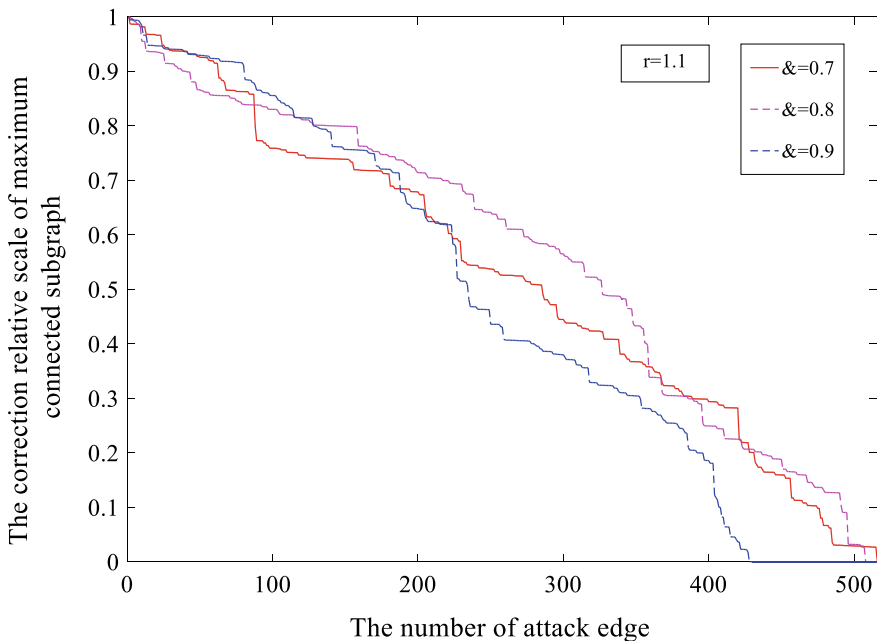


Fig. 5 Passenger flow adjustment parameter contrasts effect on invulnerability under random attacks

5 Conclusions

This paper studies the cascading failure of urban agglomeration traffic network. Firstly, this paper built the traffic network model of urban agglomeration based on the site mapping method, and proposed a new cascading failure mode based on the capacity-load model. Then we defined three node states as normal, pause and failure to give a judgment criterion based on the degree of node overload. Meanwhile, this paper mainly studied the flow redistribution mechanism of weighted network in the process of cascading transfer, and according to the overload node state, this mechanism gave the flow distribution rules which adopted the distribution based on residual capacity for pause nodes, and the distribution based on the index of connected distance weight effect for failure nodes. This paper considers the edge overload capacity between nodes and node overload capacity between edges during the cascading failure, and the results of simulation based on the actual passenger flow in Hu-Bao-E urban agglomeration are shown as follows.

- (1) The overload capacity parameter has a threshold of 1.2 with the node attacks, so this time the network invulnerability would not be improved when we increase the value of overload capacity parameter.
- (2) There is no clear correlation between initial network load and invulnerability because the large or small initial load will all reduce the network invulnerability.
- (3) Compared with the node attacks, the network has a stronger invulnerability and irregularity under the edge attacks.
- (4) The urban agglomeration traffic network has an extremely strong scale-free characteristic under the random attacks.

Our research has a stronger guiding significance for practical traffic planning than previous researches. While the disadvantage is that the dynamic characteristics of node load in the network are ignored, this means that the node or edge load is constantly changing with time. So this will be the focus of our next work.

Acknowledgements We would like to acknowledge the National Natural Science Foundation of China (No. 62063023), Natural Science Foundation of Inner Mongolia Autonomous Region (No. 2019MS05083) and Research Program of Science and Technology at Universities of Inner Mongolia Autonomous Region (No. NJZY19013).

References

1. He J, Li C, Yu Y et al (2017) Measuring urban spatial interaction in Wuhan Urban Agglomeration, Central China: a spatially explicit approach. *Sustain Cities Soc* 32:569–583
2. Gao J, Buldyrev SV, Stanley HE et al (2012) Networks formed from interdependent networks. *Nat Phys* 8(1):40
3. Gao J, Buldyrev SV, Havlin S et al (2011) Robustness of a network of networks. *Phys Rev Lett* 107(19):195701

4. Parshani R, Buldyrev SV, Havlin S (2010) Interdependent networks: reducing the coupling strength leads to a change from a first to second order percolation transition. *Phys Rev Lett* 105(4):048701
5. Buldyrev SV, Parshani R, Paul G et al (2010) Catastrophic cascade of failures in interdependent networks. *Nature* 464(7291):1025
6. Di Muro MA, La Rocca CE, Stanley HE et al (2016) Recovery of interdependent networks. *Sci Rep* 6:22834
7. Shekhtman LM, Danziger MM, Havlin S (2018) Spreading of failures in interdependent networks. In: *Diffusive spreading in nature, technology and society*. Springer, Cham, pp 397–410
8. Danziger MM, Shekhtman LM, Bashan A et al (2016) Vulnerability of interdependent networks and networks of networks. In: *Interconnected networks*. Springer, Cham, pp 79–99
9. Yin RR, Liu B, Liu HR et al (2016) Research on invulnerability of the random scale-free network against cascading failure. *Physica A Stat Mech Appl* 444:458–465
10. Zhang DX, Zhao D, Guan ZH et al (2016) Probabilistic analysis of cascade failure dynamics in complex network. *Physica A Stat Mech Appl* 461:299–309
11. Hong S, Zhang X, Zhu J et al (2016) Suppressing failure cascades in interconnected networks: considering capacity allocation pattern and load redistribution. *Mod Phys Lett B* 30(05):1650049
12. Li P, Wang BH, Sun H et al (2008) A limited resource model of fault-tolerant capability against cascading failure of complex network. *Eur Phys J B* 62(1):101–104
13. Ren J, Shen M, Mong R et al (2011) Study of reducing size of cascading failures in scale-free network. *Jisuanji Gongcheng Yu Yingyong (Comput Eng Appl)* 47(33)
14. Yuan X, Shao S, Stanley HE et al (2015) How breadth of degree distribution influences network robustness: comparing localized and random attacks. *Phys Rev E* 92(3):032122
15. Shao S, Huang X, Stanley HE et al (2015) Percolation of localized attack on complex networks. *New J Phys* 17(2):023049
16. Li S, Li L, Yang Y et al (2012) Revealing the process of edge-based-attack cascading failures. *Nonlinear Dyn* 69(3):837–845
17. Ghanbari R, Jalili M, Yu X (2017) Selective load reduction in power grids in order to minimise the effects of cascade failures. In: *Industrial electronics society, IECON 2017—43rd annual conference of the IEEE*. IEEE, pp 7481–7486
18. Ghanbari R, Jalili M, Yu X (2016) Analysis of cascaded failures in power networks using maximum flow based complex network approach. In: *Industrial electronics society, IECON 2016—42nd annual conference of the IEEE*. IEEE, pp 4928–4932
19. Fan W, Liu Z (2012) An overview on modeling of cascading failures in power grids based on complex system. *Dianli Xitong Zidonghua (Autom Electr Power Syst)* 36(16):124–131
20. Cats O, Jenelius E (2015) Planning for the unexpected: the value of reserve capacity for public transport network robustness. *Transp Res Part A Policy Pract* 81:47–61
21. Rodríguez-Núñez E, García-Palomares JC (2014) Measuring the vulnerability of public transport networks. *J Transp Geogr* 35:50–63
22. Jenelius E, Cats O (2015) The value of new public transport links for network robustness and redundancy. *Transp A Transp Sci* 11(9):819–835
23. De-Los-Santos A, Laporte G, Mesa JA et al (2012) Evaluating passenger robustness in a rail transit network. *Transp Res Part C Emerg Technol* 20(1):34–46
24. Li C, Wei L, Lu T et al (2018) Invulnerability simulation analysis of compound traffic network in urban agglomeration. *J Syst Simul* 30(02):489–496
25. Li C, Wei L, Hao Y (2016) Research on characteristics of city agglomeration compound traffic network. *J Syst Simul* 28(12):2958–2965

Impact of Intelligent Networking on Vehicles Exiting at Urban Intersections



Jun Liu and Shu-Bin Li

Abstract Intelligent network connection was expected to completely break the existing rules and orders of transportation system. The theoretical peak speed of 5G wireless bandwidth could reach more than 10 Gb per second, which enables cars to exchange information such as location, speed, and destination. It would provide a technical basis for real-time information sharing among vehicles. Based on the formal study of Neural Networks, we added Gaussian noise in the RBM top-level connecting regression, constructing a CDSHybird model that could predict traffic flow in the urban intersections. The results showed that the model proposed in this paper could reflect the movement of vehicles passing through urban intersections in various situations. Further development of interconnection would probably realize the control of traffic flow without a traffic light, which could maximize the efficiency of traffic.

Keywords Traffic flow model · 5G smart · City intersection · Urban Boltzmann machine RBM · Regressor machine supporting vector SVR

Nomenclature

X	Traffic input sequence
V	Average speed input sequence
ε	Learning efficiency
W^k	The weight matrix of the kth layer
c	Hidden layer unit offset vector of the kth layer RBM
b	The shift vector of the visible layer unit of the kth layer RBM
v_i	Visible layer node status

J. Liu

Department of Investigation, Shandong Police College, Jinan 250014, China

S.-B. Li (✉)

Department of Traffic Management Engineering, Shandong Police College, Jinan 250014, China
e-mail: li_shu_bin@163.com

h_j	Hide layer node status
v_i	Visible layer node state after reconstruction
h_j	Hidden layer node state after reconstruction
Y	Predicted traffic volume
P	Predicted value of average speed
p	Distance between cars
h	Cars' length
t_a	Straight green light duration
t_b	Left turn green light duration
Δt	Duration of yellow light
n_l	Total number of traffic lights at the intersection

1 Introduction

In 2018, China's first autonomous driving demonstration zone was established in Beijing. In August 2019, Suzhou completed the first public test road with full coverage of 5G. Intelligent connected cars have become an important direction for the transformation and upgrading of the automotive industry. Due to the delay of the 4G era and network transmission rate in the 4G era, the development of intelligent connected cars has not got breakthroughs. However, with the in-depth development of 5G technology, the key technologies of intelligent connected cars have developed rapidly [25]. Fachun [6] proposed the in-vehicle Ethernet technology, which provided a protocol framework for intelligent connected cars to form the Internet of Things. They also built an Ethernet topology through simulation and planned a blueprint for the intelligent transportation network of urban transportation. Other scholars built a traffic introduction platform for smart cars in the 5G era by building a large database as well as making full use of the "Cloud + End" technology, which fully verified that it was possible to design route for smart connected cars with the help of remote map databases [9, 11, 24].

Traffic flow prediction is an important research area of intelligent transportation systems. Traditional prediction methods include K-proximity [4], machine learning [15], and artificial neural networks [3]. These models are simple and practical, but they are not suitable for coping with huge amounts of traffic data. To solve this problem, some scholars applied Deep Learning to traffic big data. For example, Lv et al. [17] used a deep auto encoder to predict traffic flow, and Zeng and Zhang [23] proposed predicting traffic flow based on a recurrent neural network. These methods provided new and effective means for traffic flow prediction. Some Deep learning methods, such as multiple demand prediction models [5], residual network models [22], convolutional networks and recurrent network hybrid models [10], have already been used to forecast traveling demands [19]. Some scholars focus on the variability of people's daily needs of the public transportation network. Baek and Sohn [1] considered the correlation between the internal time and space of the public

transportation and propose a prediction method of DNN model. In the context of intelligent network, these traditional traffic flow models will undergo qualitative changes. Therefore, it is necessary to study the traffic flow in the intelligent network.

The purpose of intelligent connected car technology is to make the unmanned driving and autonomous driving come true, the safety needs of which is extremely high. Li et al. [12] developed an intelligent network information security framework model to provide a theoretical basis for the network security standards and research methods of the intelligent networks in the future. Wang and Yancheng [20] have verified that V2X technology can make autonomous driving of vehicles possible with electromagnetic compatibility tests. Some scholars have pointed out that control by remote sensing technology can ensure that cars maintain a safe distance from road obstacles or other vehicles [18]. The above research provides the theoretical basis for this article.

Traffic accident is one of the important factors affecting the efficiency of traffic. Moreover, traffic accidents mostly occur at urban intersections, and they greatly worsen the problems such as vehicle delays, fuel consumption, and exhaust emissions. Golob and Recker [7] studied the relationship between traffic accidents on highways and weather, lighting conditions, traffic flow and other factors based on measured traffic flow data and multivariate statistical models. Bedard et al. [2] used multiple logistic regression algorithms to study the correlation of the dying risk of driver and collisions as well as vehicle characteristics. In an intelligent networked environment, remotely controlling vehicles are bound to reduce the impact of traffic accidents. Liang et al. [14] used the smart car with four-wheel drive front wheel steering to establish a longitudinal and lateral dynamic model. Guo et al. [8] proposed predicting vehicle dynamics through big data analysis to achieve coordinated control of vehicles and coordinated vehicle–road control. Wang et al. [21] proposed a car-following model based on the Internet of Vehicles and make a stability analysis, which can reflect the specific situation of road traffic after the network-connected cars are put into use. Li et al. [13] developed a signal control algorithm to optimize vehicle trajectory in an intelligent networked environment. The results show that the throughput of intersections has improved significantly. Liu and Chang [16] proposed an optimization model for arterial road signals, which can solve the queue overflow problem at congested intersections and alleviate traffic congestion. Based on this, this article analyzes the road traffic state after the popularization of intelligent network connection technology, and explores whether the traffic light system at the intersection can be cancelled to maximize the traffic efficiency.

Because the transportation is usually characterized by large data volume and high characteristic dimension, some lower traffic volume prediction models cannot fully analyze the essential characteristics of traffic volume data changes. And the correction of the prediction can not be guaranteed either. Based on the traditional neural network, we add Gaussian noise to connect the RBM top-level regression into a CDSHybird combination model to predict traffic flow. We extract two important traffic flow parameters, traffic volume and average speed, from the traffic block port data, floating car data and Internet traffic data of Shenzhen in March 2018.

In order to study the smart traffic scheme in cross area, we make a hypothesis. In the cross area of a four-lane double way street, the traffic flow of the peak period could be 500 vehicles per hour, and two lanes have to be blocked for half an hour due to some incidents. All the vehicles are equipped with auto-driving system that can receive the management codes from the control center. The control center could work out a reasonable traffic light scheme by building a traffic flow model.

One two-way road could be blocked in three ways, left-hand lane blocked, right-hand lane blocked, and both left-hand and right-hand lanes blocked. In our model, we analyzed the time length and the distances vehicles moved. Meanwhile, we took the distances between every two vehicles and the number of vehicles into consideration when we built the model, which could be figured out after the analysis of the time and distance of the last vehicle when it passed the lane.

On the basis of this problem, this article boldly assumes that the traffic light system will be completely cancelled at the intersection, and provides design concepts for 5G intelligent transportation systems.

2 Methods and Materials

2.1 Hypothesis of Method

1. The distance between each car is large enough and the cars are equal in length.
2. Each car moves in a straight line with uniform acceleration and equals acceleration before turning.
3. Vehicles start a turning when it is perpendicular to the parking line, and end the turning when it is perpendicular to the lane.
4. When the red light is on, cars that have passed the parking line will not continue to count the time without affecting other cars.
5. When both the coming lanes are blocked, the coming cars pass the road in the leaving lane by the coming lane.
6. The distance between the first vehicle and the stop line in each lane is equal the distance between every two vehicle.
7. When the vehicle goes straight, the distance between the vehicles remains unchanged.
8. The green light time stops when the last car just turns.
9. Do not consider the impact of pedestrians and non-motor vehicles on cars and motor vehicles comply with traffic rules.

2.2 Traffic Prediction

In order to study the traffic situation at urban intersections, this paper first needs to predict the traffic flow. In order to carry out traffic prediction analysis, it is necessary

to simulate traffic congestion and regulate the collected data. This paper designs a CDSHybird combination algorithm to establish a traffic volume prediction model.

Traffic (Q), refers to the ratio of the number of vehicles passing by among the observation time (t) to the observation time (t). The number of traffic Q in t_i is x_{t_i} . The average speed is V_{t_i} , and the traffic sequence of this road section in a short period of time in the future is $x_{t_1}, x_{t_2}, \dots, x_{t_n}$, the average speed sequence is $V_{t_1}, V_{t_2}, \dots, V_{t_n}$, and each adjacent time intervals are equal.

Traffic flow prediction refers to analyzing the traffic volume data for a research section in the past period and making a prediction of the traffic volume of the section at a future time. That is, by analyzing the traffic volume data $x_{t_1}, x_{t_2}, \dots, x_{t_n}$ in the past period, we can find out the functional relationship $x_{t(n+1)} = f(x_{t_1}, x_{t_2}, \dots, x_{t_n})$ to predict a certain future moment's traffic flow. Similarly. The average speed sequence $V_{t_1}, V_{t_2}, \dots, V_{t_n}$ can be used to predict the average speed $V_{t(n+1)}$ at a certain time in the future.

Assuming that the time interval for each statistical traffic flow data is Δt . The total number of road sections studied is N . The traffic sequence of the research section would be:

$$X = (x_1, x_2, \dots, x_n) \quad (1)$$

$$V = (V_1, V_2, \dots, V_n) \quad (2)$$

Among them, $x_i = (x_{t_1}^i, x_{t_2}^i, \dots, x_{t_n}^i)$ represents the current period of the adjacent section of the studied section and the traffic volume of several historical periods, and we can find that $t_1 - t_2 = \Delta t$. The traffic flow of time i is x_{t_i} . Then the traffic volume or average speed of the next moment of any predicted road segment is predicted by the traffic flow data of the current time of the road segment and its adjacent road segments and the k previous historical moments. This article uses $X = (x_1, x_2, \dots, x_n)$ and $V = (V_1, V_2, \dots, V_n)$ as input data to predict the traffic volume or speed of the i th section. We can get the predicted traffic volume value or the predicted average speed:

$$Y = X_{tn+\Delta t}^i \quad (3)$$

$$P = V_{tn+\Delta t}^i \quad (4)$$

2.2.1 Traffic Flow Prediction Process

As to the part of traffic flow prediction, we analyzed the changes in traffic volume and average speed with time and space to determine the correlation between traffic volume and average speed with time and space. Considering the contribution of different nodes to the prediction task, we then introduced continuous CRNM, and

designed a traffic prediction model based on the CDSHybird model and an average speed prediction model that continues the CDSHybird model.

2.2.2 Spatial–Temporal Distribution of Traffic Flow

The traffic flow parameters are important factors in predicting traffic flow and judging road congestion. It changes continuously not only with time, but also with spaces. Therefore, before establishing a prediction model in this paper, we first analyze the changes in the volume and average speed of background road traffic over time.

Traffic volume refers to the number of vehicles passing through a road section in time unit. In this paper, it refers to the number of vehicles passing through a certain cross in time unit. It is an important research object in this paper to predict traffic flow. There will be some differences and similarities in its changing trends. In this paper, we study the correlation of traffic volume in different time and space by analyzing the traffic volume of the coming lanes and leaving lanes in the Municipal Party School section of Xiangmihu Road, Shenzhen City.

Traffic volume shows a certain regularity at different time periods in the Municipal Party School section and the traffic volume. Figure 1 shows the data from March 26, 2018 to March 31, 2018. From the figure, we can see that the traffic volume of each section is smallest between 2 and 6 a.m., and it grow rapidly after 7 a.m., then reaching a peak at about 9 a.m. This state can last to 9 p.m., then the traffic flow start to decline. Regularly, peaks generally occur during off hours.

Because traffic congestion on a certain section may cause congestion on the upstream section of the section, it is also necessary to study the spatial distribution of traffic volume in this article. In order to facilitate observation, we also record upstream and downstream traffic volumes of Municipal Party School section from

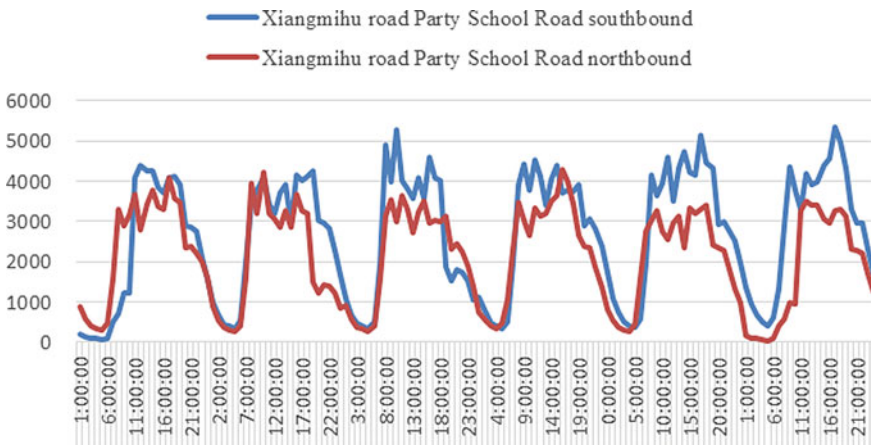


Fig. 1 Traffic volume of Xiangmihu Road Party School in the south and north

March 26 to March 31 as Figs. 2 and 3 shows. We can see that the traffic volume of the adjacent road section will decrease when it decrease in a section. When traffic congestion occurs on a road section, the traffic volume of the adjacent road section will increase in a later period; when the traffic volume of a certain road section decreases, the volume of traffic on its adjacent sections will also decrease over time.

The average speed of traffic has a lot to do with the change of traffic volume. When the traffic volume is large, the speed of the vehicle will be relatively slow. In order to betterly establish a prediction model of the average speed, we analyze the temporal and spatial characteristics of the average speed analysis. The time distribution characteristics of the section of the Party School on Xiangmihu Road, Shenzhen

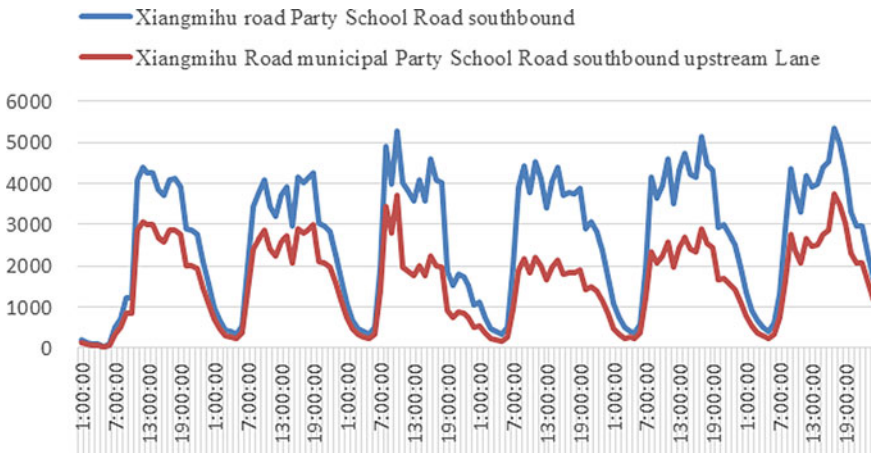


Fig. 2 Traffic volume of Xiangmihu Road Municipal Party School and its southbound section

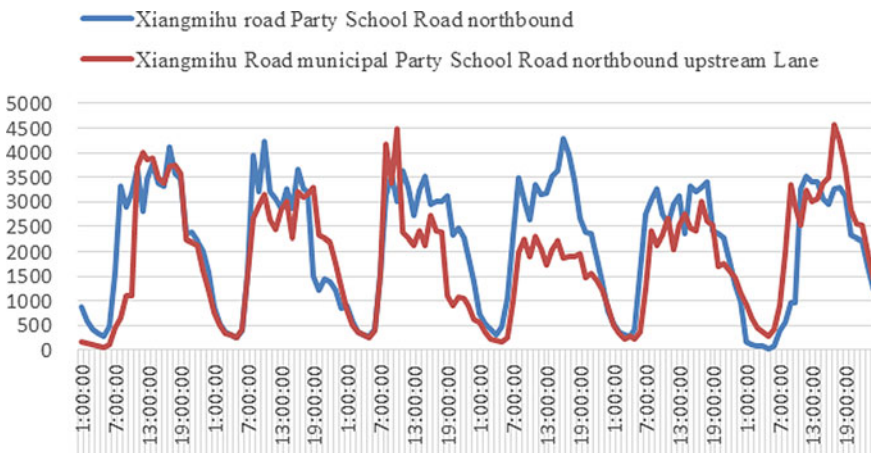


Fig. 3 Traffic volume of Xiangmihu Road Municipal Party School and its northbound section

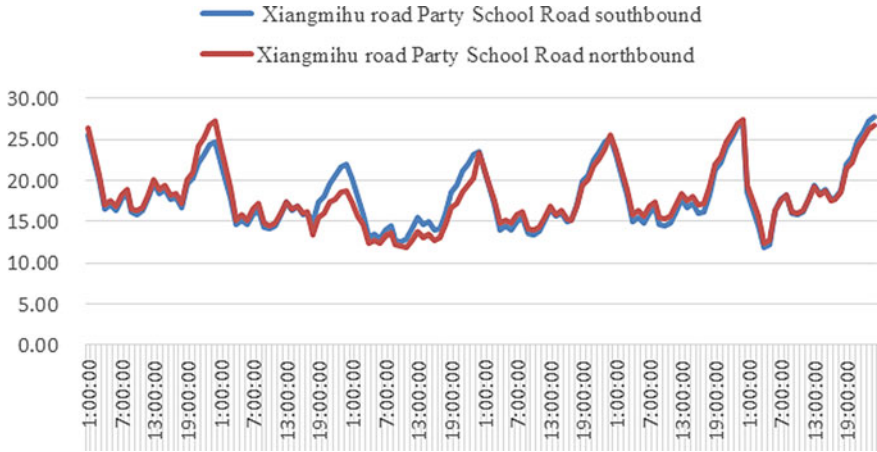


Fig. 4 Traffic average speed data of Xiangmihu Road

City are still being analyzed. The data is the average traffic speed from March 26th to March 31st, 2018, as shown in Fig. 4.

It can be seen from the figure that the daily change in the average speed of the vehicle is also showing a certain regularity. When the traffic volume increases, the average speed of the vehicle speed must decline; when the traffic volume decreases, the average speed may increase slightly. Taking the data curve of the Party School section of as an example, it can be observed that between 0 and 6 a.m., the average speed of this section is ranged from 20 to 25 km per hour. After 6 a.m., the vehicle's speed decreased rapidly. And the traffic flow also increased at this time, which is accordant to the time distribution characteristics of the traffic flow. From 10 o'clock, the speed of the vehicle slowly increases, reaching a small peak around 12 o'clock, and then start to decline, hitting the bottom at about 6 p.m., and rebound after that.

Based on the analysis of the traffic volume, we can recognize that the traffic volumes of the Party School section and the surrounding sections are different, and the average speeds are different as shown in Figs. 5 and 6.

From the figure, we can analyze the relationship between the average speed and time point. Combined with the spatial distribution characteristics of traffic volume, it is not difficult to find that the traffic flow in the upstream section of Xiangmihu road is the smallest, and the average speed of this section is the fastest. Through the above analysis, we can know that the change of traffic volume and average speed has a great relationship between the adjacent sections, so we should consider their spatial distribution when designing the prediction model of traffic volume and average speed.

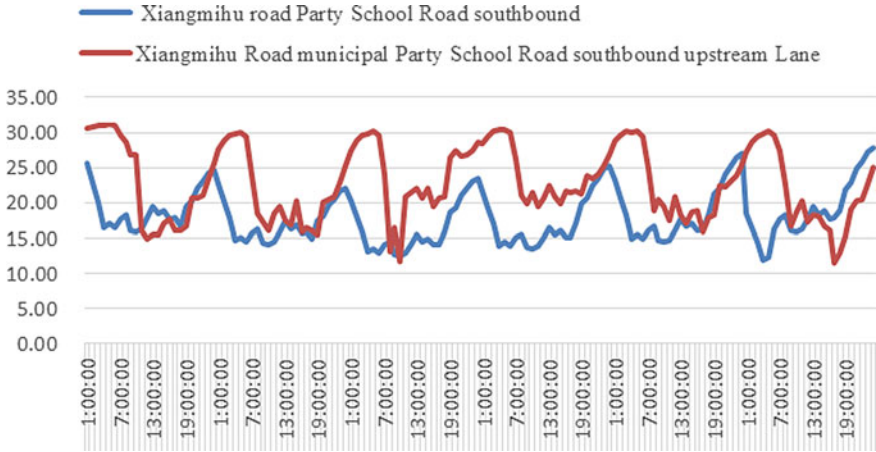


Fig. 5 Statistics of average speed of Xiangmihu Road Municipal Party School and its southbound section

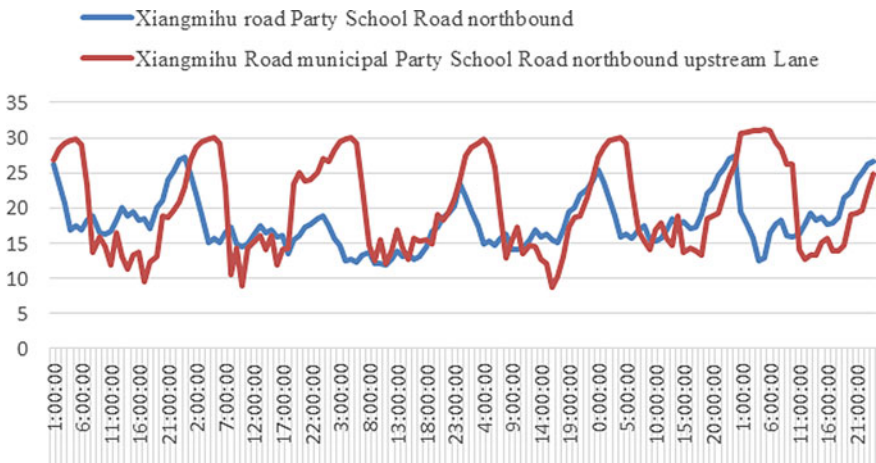


Fig. 6 Statistics of average speed of Xiangmihu Road Municipal Party School and its northbound section

2.2.3 Traffic Flow Prediction Model Based on CDSHybird Model

After analyzing the temporal and spatial correlation between traffic volume and average speed, we tried to build a new model, which involved the contribution of different nodes to the prediction task. We added Gaussian noise to make the nodes continuously, and connected the regression model at the top to make global adjustment. Then we built the CDSHybird model to predict traffic volume.

RBM is a formula based on energy. It assumes that the number of hidden layers is m , and the number of visible layers is n . When the state of (v, h) is determined, the energy formula of RBM is:

$$E(v, h|\theta) = -\sum_{i=1}^n b_i v_i - \sum_{j=1}^m c_j h_j - \sum_{i=1}^n \sum_{j=1}^m c_i w_{ij} h_j \quad (5)$$

Among them, $\theta = \{w_{ij}, a_i, b_j\}$ are parameters of RBM; V_i and h_j are visible layer neurons and hidden layer neurons, b_i and c_i are visible layer bias and hidden layer bias respectively, w_{ij} is connection weight between visible layer and hidden layer. From the energy formula of RBM, it can be concluded that the joint probability distribution of (v, h) :

$$P(v, h|\theta) = \frac{e^{-E(v, h|\theta)}}{z^\theta} \quad (6)$$

$$z(\theta) = \sum_{v, h} e^{-E(v, h|\theta)} \quad (7)$$

where $z(\theta)$ is the normalization factor. The distribution $P(v|\theta)$ of v can be obtained. And the marginal distribution of the joint probability distribution $P(v, h|\theta)$ is:

$$P(v|\theta) = \frac{1}{Z(\theta)} \sum_h e^{-E(v, h|\theta)} \quad (8)$$

Considering the structural characteristics of RBM, we can get the conclusion that if the visible unit (or hidden unit) is known, and the activation states of each hidden unit (or visible unit) are independent from each other, then the activation probabilities of the j th hidden unit and the i th visible unit are respectively:

$$P(h_j = 1|v, \theta) = \sigma\left(c_j + \sum_i v_i w_{ij}\right) \quad (9)$$

$$P(v_i = 1|h, \theta) = \sigma\left(b_i + \sum_j h_j w_{ij}\right) \quad (10)$$

$\sigma(x) = \frac{1}{1+\exp(-x)}$ is the sigmoid function.

If v_i is known, then the conditions of h_j nodes will be independent. Similarly, if h_j is known, then the conditions of v_i nodes will be independent. There are:

$$P(h|v) = \prod_J P(h_j|v) \quad (11)$$

$$P(v|h) = \prod_J P(v_j|h) \quad (12)$$

Thus, if the visible layer node is known, we can calculate the probability of activating the j th hidden layer node. Similarly, when the hidden layer node is known, we can calculate the probability for activating the i th visible layer node.

$$P(h_j = 1|v) = \frac{1}{1 + \exp(-\sum_i v_i w_{ij} - c_j)} \quad (13)$$

$$P(v_i = 1|h) = \frac{1}{1 + \exp(-\sum_j h_j w_{ij} - b_i)} \quad (14)$$

RBM is a probability model. Compared with other neural networks, the probability generation model is used to establish the distribution relationship between the observed data and its label. But a single RBM has some shortage as it is a neural network composed of a visible layer and an output layer. It can only get shallow features, instead of deep features in data feature analyzing. Since the support vector regression model has a good effect on the processing of nonlinear data, so we consider to connect a SVR predictor at the top of RBM to predict traffic volume.

Support Vector Regression (SVR) is a regression algorithm developed on the basis of Support Vector Machine (SVM), which is mainly used to solve regression problems. It is similar to SVM, but there is only one sample point of SVR. The optimal hyperplane is not to maximize the separation of the two types of sample points, but to minimize the "Total Deviation" of all sample points from the hyperplane. At this time, the sample points are all between two boundary lines, so the optimal regression hyperplane is the same as the maximum interval.

In a set of training samples $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $x_i \in R^m$ is the input data, and $y_i \in R$ is the output data. Classification learning will create a for multilayer hyperplane based on the training set in the sample space. Standard support vector machine uses $\varepsilon - Insensitivity$ function, which assumes that all training data are fitted with linear function under the accuracy ε . Support vector regression can only be used when $f(x)$ and y only has ε deviation. Only when the absolute value between $f(x)$ and y is greater than ε , the loss value can be calculated.

Firstly, nonlinear mapping is used to map the original input data from the original space to the N-dimensional feature space, and the linear regression calculation is completed in the high-dimensional space. So the function is constructed in the feature space. In this way, the nonlinear regression problem in low dimensional space can be transformed into linear regression problem in high dimensional space. The loss function can be defined as following:

$$L_\varepsilon(x_i, y_i, f) = \max\{0, |y_i - f(x_i)| - \varepsilon\} \quad (15)$$

We need to find a suitable function to minimize $H(w)$:

$$H(w) = \frac{1}{2}(w \cdot w) + C \frac{1}{m} \sum_{i=1}^m L_\varepsilon(x_i, y_i, f) \tag{16}$$

Among them, $w = (w_1, w_2, \dots, w_N)^T$ is the weight number, and C is the adjustable parameter. By introducing the nonnegative variables ξ_i and ξ'_i , the following results are obtained:

$$\min \frac{1}{2}(w, w) + C \sum_{i=1}^m (\xi_i + \xi'_i) \tag{17}$$

The constraints are:

$$\begin{cases} [w \cdot \phi(x_i) + b] - y_i \leq \varepsilon + \xi_i \\ y_i - [w \cdot \phi(x_i) + b] \leq \varepsilon + \xi'_i \\ \xi_i, \xi'_i \geq 0, i = 1, 2, \dots, m \end{cases} \tag{18}$$

The corresponding dual problems are as follows:

$$\begin{cases} \max \sum_{i=1}^m y_i(\beta_i - \beta'_i) - \varepsilon \sum_{i=1}^m (\beta_i + \beta'_i) - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\beta_i - \beta'_i)(\beta_j - \beta'_j) K(x_i, y_j) \\ s.t. \sum_{i=1}^m (\beta_i - \beta'_i) = 0, 0 \leq \beta_i, \beta'_i \leq C/m \end{cases} \tag{19}$$

Among them, β_i and β'_i are Lagrangian coefficients of ξ_i and ξ'_i respectively, and $K(x_i, y_j)$ is its kernel function.

To sum up, we can get the corresponding output of the test sample as follows:

$$f(x) = \sum_{i=1}^m (\beta_i - \beta'_i) K(x_i, y_j) + b \tag{20}$$

By the traffic volume and average speed data from March 26 to March 30, we can predict the traffic volume and average speed of Xiangmihu Road southbound on March 31. The comparison between the predicted value and the actual value is as in Figs. 7 and 8.

Both figures are based on the information of the southbound section of Xiangmihu road from March 26 to March 30. It is found that the information prediction of the average speed is more accurate. There is a large deviation in the prediction of traffic flow between 9:00 and 16:00. The reason may be that the sample data volume is too small.

To sum up, we can see that the traffic volume prediction model based on CDSHy-bird designed in this paper has a good effect on the fitting of the predicted value

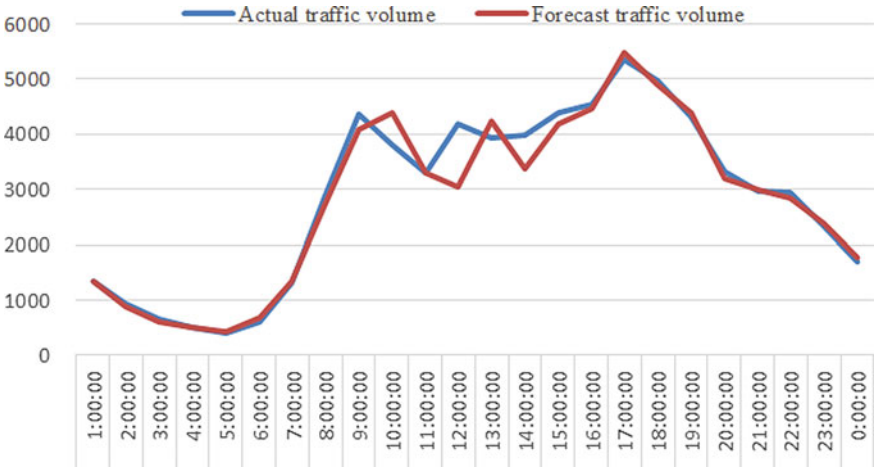


Fig. 7 Comparison between the actual traffic volume and the predicted traffic volume of Xiangmihu Road

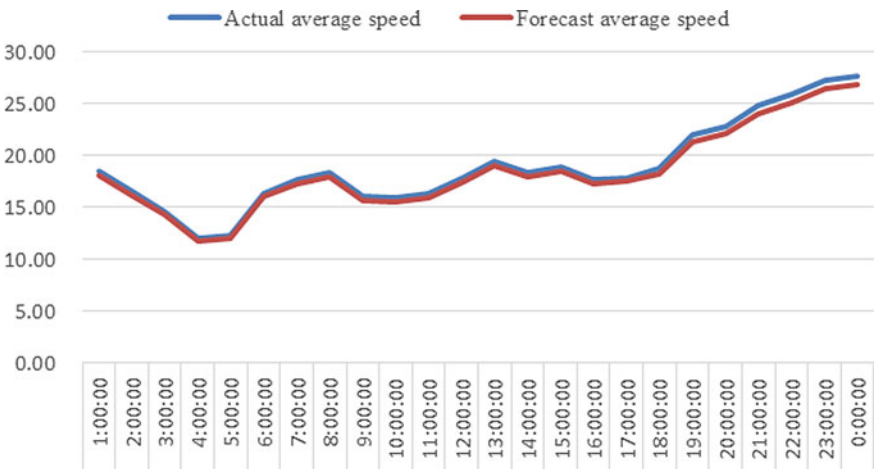


Fig. 8 Comparison between the actual and predicted average speed of the southbound section of Xiangmihu Road

of traffic volume and average speed with the real data, and the prediction effect is slightly worse in the local area with large traffic flow fluctuation.

2.3 Method Description

2.3.1 Characterization of Vehicle Road Occupancy in Extreme Cases

Road occupancy is the ratio of the distance vehicles take to the length of the lane, and is shown as β_i . We can get the formula as below:

$$\beta_i = \frac{(p+h)Y}{d_i} \quad (21)$$

We use $T_{GE}, T_{GW}, T_{GN}, T_{GS}, T_{RE}, T_{RW}, T_{RN}, T_{RS}$ to indicate one Duration of the green light in the east of the intersection, the green light in the west, the green light in the north, the green light in the south, the red light in the east, the red light in the west, and the red light in the north, and the red light in the south. When the traffic light at a certain position of the intersection is displayed as a green light, the traffic light at the opposite position must be displayed as a green light, and the duration of the green lights in both directions is equal. Then the number of vehicles that pass the cross in one green light can be expressed as:

$$N = 2 \left[\frac{T_{GE} \cdot v}{p+h} \right] \quad (22)$$

In order to get the total number of vehicles passing through the intersection during the green light time, we need to double N . The ultimate non-traffic state is that during the green light time, all vehicles on the road pass the stop line just before the right light is on. The road occupancy rate of the ultimate non-traffic state is represented by β_0 , which can be obtained:

$$\beta_0 = \frac{2p+h \cdot \left[\frac{T_{GE} \cdot v}{p+h} \right]}{d_i} \quad (23)$$

Among them, d_i represents the length of the i th road.

2.3.2 Characterization of Vehicle Scheduling Process

When it is necessary to close the two lanes g_3 and g_4 due to traffic accidents, road maintenance, etc., we will use the b_4 lane to replace the two lanes g_3 and g_4 (Fig. 9).

When the car turns, the car's motion trajectory is exactly close to a quarter of an ellipse. Therefore, in the calculation, we use the algorithm to calculate the ellipse motion trajectory to quantify the various parameters of the car turning. This article positions the end point of lane b_4 from west to east to the midpoint of lanes b_7 and b_8 the end point of lane of b_4 lane from west to north to the midpoint of lanes b_5 and b_6 , and the lane b_4 from west to south to the midpoint of lanes b_1 and b_2 . This

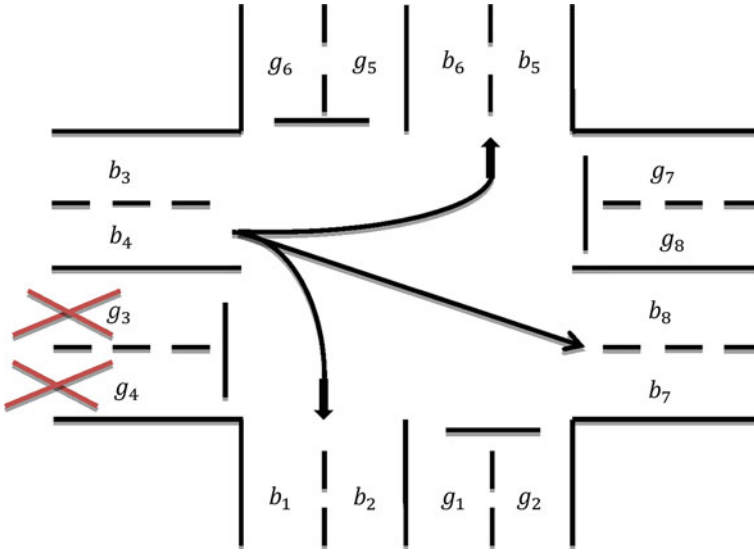


Fig. 9 Schematic diagram of two driving out lanes closed

can simplify the research of the problem. From the algorithm of the elliptical motion trajectory, we can get the following formula:

$$l = 4a \int_0^{\frac{\pi}{2}} \sqrt{1 - e^2 \sin^2 \beta} d\beta \approx 2\pi b + 4(a - b) \tag{24}$$

Among them, a represents the semi-major axis and b represents the semi-minor axis of the ellipse of the car. Assuming that the width is equal to x , the vehicle turning trajectory coefficient can be obtained (Table 1).

When the car goes straight from west to east from lane b_4 , the car runs roughly diagonally to the right. We assume that the car’s motion is a uniformly accelerating linear motion when going straight. When the green light on east–west direction is on, the vehicles waiting for the green light in lane b_4 will have two movement trajectories, going straight from west to east or turning from west to south. At this time, vehicles

Table 1 Vehicle turning trajectory coefficient

	From b_4 to b_1, b_2	From b_4 to b_5, b_6	From g_1 to b_3	From g_1 to b_7, b_8	From g_5 to b_3	From g_5 to b_7, b_8	From g_7 to b_1, b_2	From g_7 to b_5, b_6
a	$\frac{5}{2}x$	$3x$	$\frac{7}{2}x$	$\frac{3}{2}x$	$\frac{3}{2}x$	$3x$	$3x$	x
b	x	$\frac{3}{2}x$	$\frac{5}{2}x$	x	$\frac{1}{2}x$	$\frac{5}{2}x$	$\frac{5}{2}x$	$\frac{1}{2}x$

waiting for a green light need to be analyzed. This article discusses two extreme situations, that is, all waiting vehicles in lane b_4 turn right and all go straight.

We took all the vehicles in a queue as a train fleet, and the maximum speed allowed for vehicles to pass the intersection as the highest speed of the fleet. Teams waiting for the green light are considered to be accelerating after launch. We weakened the influence of the centripetal force of the circular motion on the acceleration as much as possible, and compared the overall trajectory lengths of the two to find that the shorter the time it took when the vehicle turned to the right, the bigger the amount of the cars passed.

We assume that the time when traffic lights go straight from the north to the north as the starting point, and take 250 vehicles as the research object, distributing them evenly in the four directions of the intersection. Then, each direction of the crossroads is divided into 63 vehicles. According to this relationship, the following functional relations are listed:

$$\text{ceil}\left(\frac{vt_a}{p+h}\right)n_l \geq 63 * k \quad (25)$$

$$\text{ceil}\left(\frac{vt_b}{p+h}\right)n_l \geq 63 * (1 - k) \quad (26)$$

Due to the blockage of lanes g_3 and g_4 , vehicles in lanes g_1 and g_2 will not be affected by turning from the left nor going straight, and will not be affected by lanes g_3 and g_4 . In order to assure all vehicles to pass through the intersection quickly, we need to schedule these vehicles so that they can pass through the intersection no matter the lights go straight is on or the lights turn left is on. So the following functional relations are listed:

$$\text{ceil}\left(\frac{v(t_a + t_b)}{p+h}\right)2n_l \geq 84 \quad (27)$$

When two lanes b_1 and b_2 need to be closed due to traffic accidents, road maintenance, etc. (Fig. 10).

When lanes b_1 and b_2 are blocked, vehicles in lane g_7 cannot perform turn left. In case of traffic congestions, we need to rearrange them. That is, they will go straight along g_8 . Then, we can get the following formula:

$$\text{ceil}\left(\frac{vt_a}{p+h}\right)2n_l \geq 63 \quad (28)$$

When one of the two closed lanes is leaving the lane and one is entering the lane, the calculation method is the same as the previous one (Fig. 11).

And formulate the objective function for the above three cases:

$$T_{\min} = n(t_a + t_b + 4\Delta t) \quad (29)$$

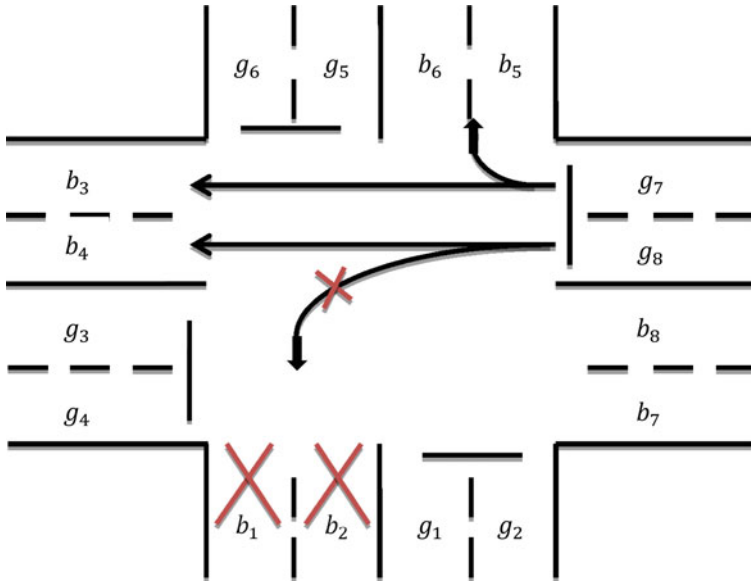


Fig. 10 Schematic diagram of two driving in lanes closed

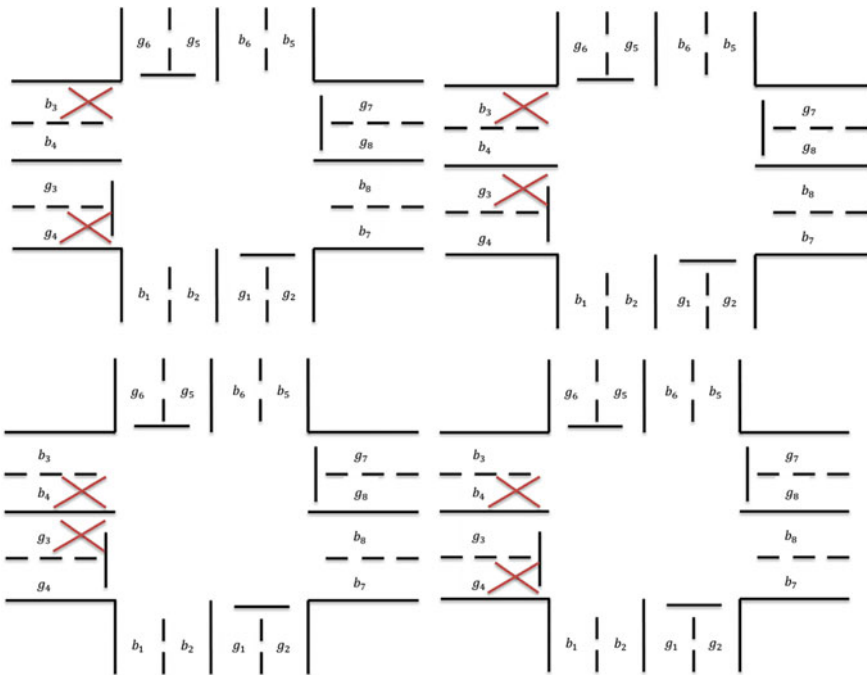


Fig. 11 Other road closure situations

2.4 Results

We use the Lingo11 software to form a non-linear programming equations and solve them. Then, we get the appendix.

The result obtained is:

- (1) $t_a = 37.83938$; $t_b = 39.06$; $n_1 = 1$; $k = 0.4920635$; $T_{\min} = 84.89938$
- (2) $t_a = 37.83938$; $t_b = 39.06$; $n_1 = 1$; $k = 0.4920635$; $T_{\min} = 84.89938$
- (3) $t_a = 51.66$; $t_b = 51.66$; $n_1 = 1$; $k = 0.5$; $T_{\min} = 111.32$.

2.5 Traffic Light Control Scheme and Route Arrangement

After analyzing the above three types of situations, we find that when the proportions of the vehicles going straight forward and right-turning are almost the same, then the traffic efficiency is highest. That means the total time the vehicles take to pass the cross thoroughly is shortest within the normal endurance of the driver.

- (1) The duration of the green light of the straight traffic signal is 37.84 s, and the duration of the left turn signal is 39.06 s. The number of vehicles on the straight road accounts for 49.20635% of the total number of vehicles, and the total time for all vehicles to pass is 84.90 s.
- (2) The duration of the green light of the straight traffic signal is 37.84 s, and the duration of the left-turn light is 39.06 s. The number of vehicles on the straight road accounts for 49.20635% of the total number of vehicles, and the total time for all vehicles to pass is 84.90 (s).
- (3) The duration of the green light of the straight traffic signal is 51.66 s. The duration of the left turn traffic signal is 51.66 s, and the traffic on the straight road accounts for 50% of the total number of vehicles.

The route adjustment plan based on the solution worked out as follows:

1. The first type of situation: The cars in lanes $g_1, g_2, g_3, g_4, g_5, g_6$ and g_8 pass normally. The vehicles in lane g_7 need to be dispatched to go straight with the cars in lane g_8 .
2. The second type of situation:
 - (1) Cars in lanes g_2, g_3, g_5, g_6, g_7 and g_8 pass normally. Vehicles in lane g_4 can turn right into lane b_2 , and vehicles in lane g_6 can go straight when the left-turn light is on.
 - (2) Cars in lanes g_1, g_3, g_5, g_6, g_7 and g_8 pass normally. Vehicles in lane g_4 can turn right into lane b_2 . Vehicles in lane g_5 can turn left when the lights go straight across is on.
 - (3) Cars in lanes g_2, g_3, g_4, g_5, g_6 and g_8 pass normally. The vehicles in lane g_7 need to be dispatched to go straight with the cars in lane g_8 .
 - (4) Cars in lanes g_1, g_3, g_4, g_5, g_6 and g_8 pass normally. Vehicles in lane g_7 need to be dispatched to go straight with cars in lane g_8 .

- 3. The third type of situation: The cars in lanes g_3, g_4, g_5, g_6, g_8 and g_8 pass normally. Originally, vehicles from the g_1 and g_2 directions are dispatched to the other three directions.

3 Simulation

In our research process, we mainly talked about the situations that some lanes are closed in the intersection. However, the traffic system is a holistic system. Any traffic accident or road construction may affect the overall traffic situation, and there is a certain functional relationship between the range and the size of the influence. By studying the two crossroads directly related to the closed lane and the six crossroads that are indirectly related to the closed lanes, we transformed the partial problems into global problems. The impact of a closed road on the global situation can be roughly expressed by using a three-dimensional peak wave diagram (Figs. 12 and 13).

In Fig. 12, the plane formed by x and y is a horizontal plane, the unit of x and y is kilometer, and the z -axis is a simulated traffic pressure value, which means the numbers of per minute. When a section of road is closed, the two intersections closest to the incident will be the first to be affected, and the affection radiated outward in

Fig. 12 Simulation of traffic pressure

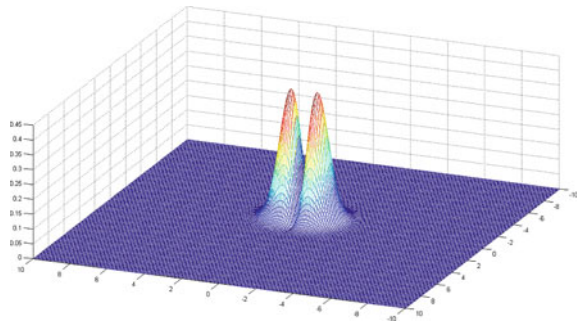
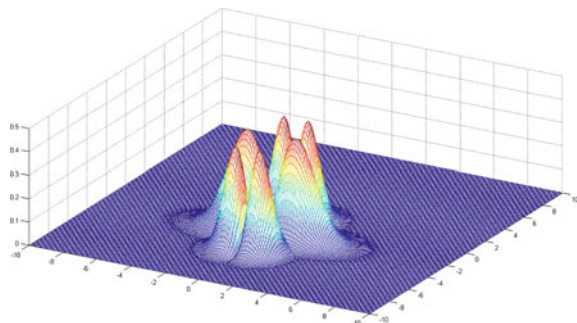


Fig. 13 Umbrella radiation pattern of traffic pressure



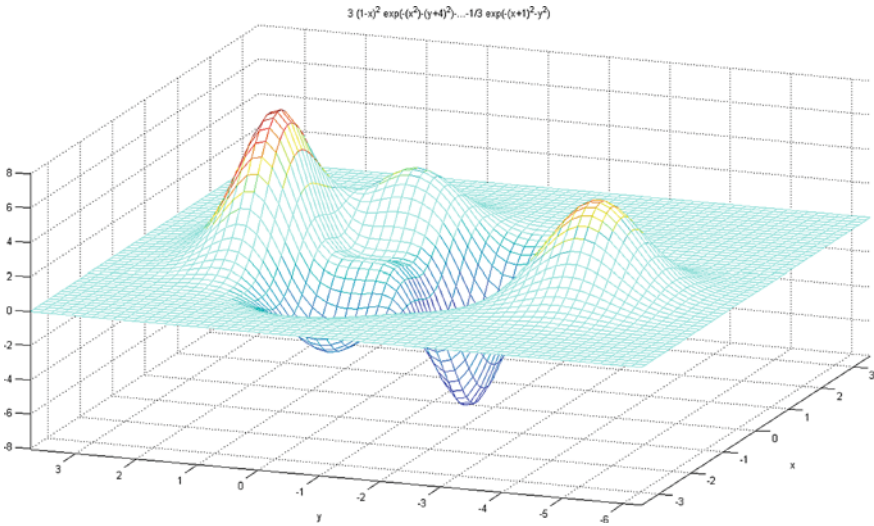


Fig. 14 Possible traffic pressure situation in the city at a certain moment

a wave. The six intersections infected by the first two sections will suffer the most affection of the block down of the lanes.

We conjecture that the degree of impact will be impacted by the distance between the tragic spot and the lane. We make the congestion at the two intersections on the ends of the road where the situation occurs two peaks. The pressure radiation can form an umbrella diagram. The extent of the impact and the distance from the incident point roughly correspond to the peak system model (Fig. 14).

The disparity of traffic pressure on both sides of the intersection at the same time may be huge, which makes the road utilization rate is relatively low. The Peaks model can be used to simulate the possible traffic pressure situation in the city at a certain moment. It can be seen that there is a large imbalance in the traffic pressure in the city.

It can be seen from Fig. 15 that the traffic pressure situation at urban intersections has its regularity, but unexpected accident traffics will break this law. In order to alleviate this situation, an intelligent control system can be set for each intersection to minimize the impact of closed roads on the traffic system. An overall intelligent control system can be established to balance traffic through 5G remote control.

4 Conclusions

Through the above research, it can be found that the traffic pressure at each intersection when a sudden accident happened on the road can be quickly relieved by the control and adjustment of the intelligent network connection system. So we have a

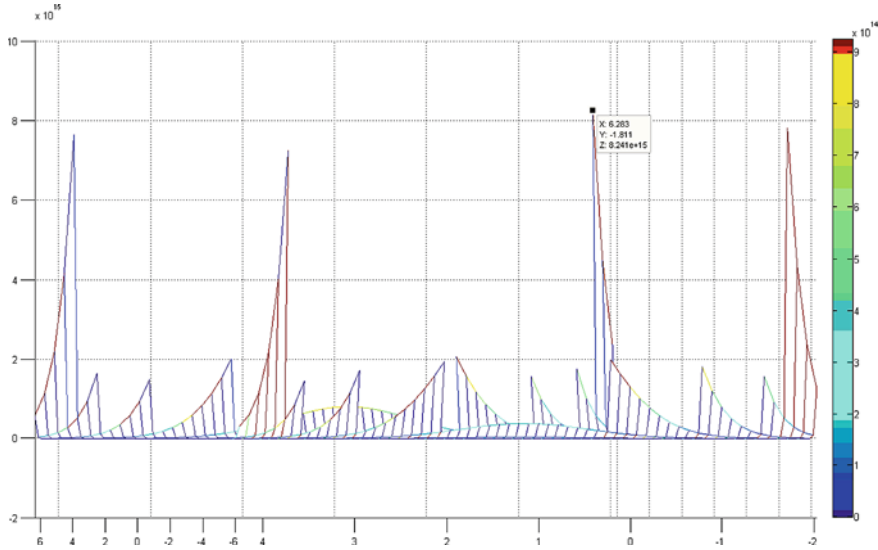


Fig. 15 Simulation of traffic pressure at urban intersections

bold outlook on whether we can control the vehicles through 5G remote control so that the distance between the cars from all sides can at least be perfectly staggered when passing through the intersection. Then we can achieve the idea of eliminating traffic lights at crossroads and maximizing the efficiency of the transportation system.

With the model established earlier, and the following questions should be discussed.

- (1) When the amount of the vehicles is small, they could pass through the intersection automatically under the control of intelligent system instead of stopping at the red light if there is a traffic light. Thus the efficiency of traffic will obviously improve.
- (2) When there are many vehicles, and no traffic light, a control system can make cars start and stop by the control of auto-driving system, so that more vehicles will not be affected by complicated external environment. Which means no traffic accidents such as a collision or a rampant will affect the traffic.

Because passengers have different destinations and urgency, it is necessary to design a traffic scheme for each vehicle separately which could form an intelligent transportation system. When there is a limit speed for the cars, they can go forward with different speed. That is, vehicles are allowed to overtake. So, we need to take every vehicle that will pass through the intersection into account. We need intelligent systems to make intelligent decisions through big data analysis and cloud computing. In the future, when people research intelligent cars, the following basic Principles should be obeyed:

- (1) Each vehicle is driven to its destination without stopping by the control system.

- (2) Vehicles are not allowed to run at the same speed, and they can do variable-speed movement.
- (3) All vehicles are unmanned and remotely controlled by intelligent systems.

Acknowledgements This work is supported by the National Natural Science Foundation of China (Grant Nos. 71871130, 71471104, 71771019 and 71971125); the University Science and Technology Program Funding Projects of Shandong Province (Grant No. J17KA211); the Open Project of Key Laboratory of Automotive Engineering of Sichuan Province (Grant No. szjj2017-014).

References

1. Baek J, Sohn K (2016) Deep-learning architectures to forecast bus ridership at the stop and stop-to-stop levels for dense and crowded bus networks. *Appl Artif Intell* 30(9):861–885. <https://doi.org/10.1080/08839514.2016.1277291>
2. Bedard M, Guyatt GH, Stones MJ, Hirdes JP (2002) The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. *Accid Anal Prev*. <https://doi.org/10.1093/geront/41.6.751>
3. Chan KY (2012) Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg-Marquardt algorithm. *IEEE Trans Intell Transp Syst* 13(2):644–654. <https://doi.org/10.1109/tits.2011.2174051>
4. Chang H, Lee Y, Yoon B, Baek S (2012) Dynamic near-term traffic flow prediction: system-oriented approach based on past experiences. *IET Intell Transp Syst* 6(3):292. <https://doi.org/10.1049/iet-its.2011.0123>
5. Chen C, Zhang D, Li N, Zhou ZH (2014) B-planner: planning bidirectional night bus routes using large-scale taxi GPS traces. *IEEE Trans Intell Transp Syst* 15(4):1451–1465. <https://doi.org/10.1109/TITS.2014.2298892>
6. Fachun L (2013) Research on the development and application of intelligent network teaching system. In: *Intelligence computation and evolutionary computation*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-31656-2_121
7. Golob TF, Recker WW (2003) Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions. *J Transp Eng* 129(4):342–353. [https://doi.org/10.1061/\(asce\)0733-947x\(2003\)129:4\(342\)](https://doi.org/10.1061/(asce)0733-947x(2003)129:4(342))
8. Guo G, Xu YG, Xu T, Li DD, Wang YP, Yuan W (2019) Overview of the Internet-sharing shared-road-coordinated intelligent transportation system. *Control Decis* 34(11):2375–2389. <https://doi.org/10.13195/j.kzyjc.2019.1316>
9. Ismaila AK, Sunday OO (2019) A big data anonymous batch verification scheme with conditional privacy preservation for power injection over vehicular network and 5G smart grid slice. *Sustain Energy Grids Netw* 20. <https://doi.org/10.1016/j.segan.2019.100260>
10. Ke J, Zheng H, Yang H (2017) Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach. *Transp Res Part C Emerg Technol* 85:591–608. <https://doi.org/10.1016/j.trc.2017.10.016>
11. Kim J (2016) The platform business model and business ecosystem: quality management and revenue structures. *Eur Plann Stud* 1–20. <https://doi.org/10.1080/09654313.2016.1251882>
12. Li Y, Cao Y, Qiu H, Gao L, Du Z, Chen S (2016) Big wave of the intelligent connected vehicles. *China Commun* 13(Supplement2):27–41. <https://doi.org/10.1109/CC.2016.7833458>
13. Li Z, Elefteriadou L, Ranka S (2014) Signal control optimization for automated vehicles at isolated signalized intersections. *Transp Res* 49c:1–18. <https://doi.org/10.1016/j.trc.2014.10.001>

14. Liang YX, Li YN, Yu YH, Zheng L (2019) Vertical and horizontal decoupling control of smart car based on neural network inverse system. *J Hunan Univ* (10):26–35. <https://doi.org/10.16339/j.cnki.hdxzbzkb.2019.10.004>
15. Lippi M, Bertini M, Frasconi P (2013) Short-term traffic flow forecasting: an experimental comparison of time-series analysis and supervised learning. *IEEE Trans Intell Transp Syst* 14(2):871–882. <https://doi.org/10.1109/TITS.2013.2247040>
16. Liu Y, Chang GL (2011) An arterial signal optimization model for intersections experiencing queue spillback and lane blockage. *Transp Res* 19C(1):130–144. <https://doi.org/10.1016/j.trc.2010.04.005>
17. Lv Y, Duan Y, Kang W, Li Z, Wang FY (2015) Traffic flow prediction with big data: a deep learning approach. *IEEE Trans Intell Transp Syst* 16(2):865–873. <https://doi.org/10.1109/TITS.2014.2345663>
18. Qing L, Tao H, Guodong F (2019) Judgment and optimization of video image recognition in obstacle detection in intelligent vehicle. *Mech Syst Signal Process*. <https://doi.org/10.1016/j.ymsp.2019.106406>
19. Vlahogianni EI, Karlaftis MG, Golias JC (2014) Short-term traffic forecasting: where we are and where we are going. *Transp Res Part C Emerg Technol* 43(1). <https://doi.org/10.1016/j.trc.2014.01.005>
20. Wang J, Yancheng LI (2006) Dynamic simulation and test verification of MR shock absorber under impact load. *J Intell Mater Syst Struct* 17(4):309–314. https://doi.org/10.1142/9789812702197_0120
21. Wang YQ, Zhou CF, Li WK, Yan BW, Jia B, Wang JX (2018) Stability analysis and wave dynamics of an extended hybrid traffic flow model. *Mod Phys Lett B* 1850055. <https://doi.org/10.1142/S0217984918500550>
22. Xu J, Rahmatizadeh R, Bölöni L, Turgut D (2018) Real-time prediction of taxi demand using recurrent neural networks. *IEEE Trans Intell Transp Syst* 19(8):2572–2581. <https://doi.org/10.1109/TITS.2017.2755684>
23. Zeng X, Zhang Y (2013) Development of recurrent neural network considering temporal-spatial input dynamics for freeway travel time modeling. *Comput-Aided Civ Infrastruct Eng* 28(5):359–371. <https://doi.org/10.1111/mice.12000>
24. Zhang J, Wang T, Li B (2014) Research on technical standardization for electric demand response. *Proc CSEE* 34(22):3623–3629. <https://doi.org/10.13334/j.0258-8013.pcsee.2014.22.006>
25. Zhao KX, Fan WW, Cui ZH, Yuan X, Chang C (2019) Exploration on talent cultivation of automotive electronic technology specialty under the background of intelligent connected vehicle. *J Henan Mech Electr Eng Coll* 027(003):70–72. <https://doi.org/10.3969/j.issn.1008-2093.2019.03.017>

Safety Evaluation of Expressway Linear Operation Based on Design Consistency



Ke Chen, Jian-Feng Jiang, Jia-Jia Tian, Yun-You Liu, and Yanli Ma

Abstract In order to solve the problem of coordination between speed and alignment design of expressway in China, and improve the safety of expressway operation. On the basis of determining the operation safety evaluation index, combined with the existing domestic traffic safety evaluation research data, referring to the American interactive highway safety design model (IHSDM), this paper build an improved highway alignment operation safety evaluation model considering design consistency. A case study is carried out to verify the effectiveness of the model. The results show that the accuracy of the model is 83.33%, which can effectively evaluate the operation safety of highway alignment design, and the research results are positive. It is of great theoretical and practical significance to guide the safety evaluation of highway alignment operation in China.

Keywords Traffic engineering · Design consistency · Linear design · Evaluation model · Operational safety

1 Introduction

With the rapid development of highway construction in China, while providing efficient and convenient services for transportation, expressways also have hidden safety problems, which are reflected in the lack of coordination between speed and highway alignment design. It is very important to carry out linear highway operation safety evaluation considering design consistency to improve highway operation safety.

The relevant research on highway operation safety evaluation abroad is relatively earlier than that in China. In 1988, the Highway Safety Audit Guide was compiled in the United Kingdom. In 1994, Australia issued the Highway Safety Assessment.

K. Chen · J.-F. Jiang · Y.-Y. Liu
Heilongjiang Highway Survey and Design Institute, Harbin 150080, China

J.-J. Tian · Y. Ma (✉)
School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China
e-mail: mayanli@hit.edu.cn

In 2002, Germany published the Road Safety Audit Recommendation. In 2018, the Federal Highway Administration released the latest version of the Highway Safety Audit Guide and the Interactive Highway Safety Design Model (IHSDM). The research on road safety evaluation in China is relatively late compared with abroad. In 1998, Tongji University developed a highway safety evaluation procedure for high-grade highways. In 2004, the department of transportation published the Highway Safety Audit (HSA). In 2015, the HAS were revised, and released as the Specifications for Highway Safety Audit (SHSA) [1–19].

Based on the existing traffic safety evaluation research in China, combined with the SHSA and referring to the IHSDM, this paper proposes a traffic safety evaluation method suitable for expressway alignment design in China, in order to solve the problem of coordination between highway operating speed and alignment design in China. Firstly, the required data and operational safety evaluation indexes are determined. Then, the expressway operational safety evaluation model based on design consistency is established. Finally, the effectiveness of the model is verified through case analysis.

2 Evaluation Indexes and Data Requirements

2.1 Evaluation Indexes

When the vehicle is driving on the highway and the highway alignment changes, the driver will immediately take measures to adjust the operating speed to ensure driving safety. Therefore, highway alignment is one of the important indexes for the safety evaluation.

The change of operating speed is directly related to the safety of road operation. The greater the speed difference between adjacent sections, the worse the safety of road operation. The transition section between tangent segment and curve segment is the most critical position for road safety evaluation. Therefore, the difference of operating speed between tangent section and curve section is an important index to evaluate the design consistency.

In the process of driving, the operating speed of vehicles is affected by many factors, and it is difficult to keep consistent with the design speed (V_{design}). Especially in the case of low design speed, the operating speed is likely to exceed the design speed. The coordination between the operating speed and the design speed is another important factor for the safety evaluation of road operation. Therefore, the difference between the operating speed and the design speed in the same section is another important index to evaluate the design consistency.

Table 1 Design consistency evaluation data requirements

Data requirements	Data element
Horizontal alignment data	Horizontal tangents, transition curves, horizontal curves
Vertical alignment data	Longitudinal slopes, vertical curves
Speed data	Design speed, limited speed, desired speed

Table 2 Desired speed requirements at different design speeds

Design speed (km/h)	100 or 120	80
Desired speed (km/h)	120	110

2.2 Data Requirements

The data required to evaluate the design consistency include the highway alignment data and speed data, as shown in Table 1.

Horizontal alignment data mainly include data of tangent segment, transition curve segment and circular curve segment; The data of tangent segment include the starting and ending pile numbers; Data of transition circular curve segment and circular curve section include starting and ending pile numbers and radius of circular curve. Vertical alignment data mainly include data of longitudinal slope segment and vertical curve segment; The data of vertical curve segment include the starting and ending pile numbers, radius of vertical curve and tangent length of the curve.

Speed data include design speed, limit speed, and desired speed. The design speed refers to the maximum operating speed that general drivers can maintain safely and comfortably when the climate is normal, the traffic density is small, and the vehicle operation is only affected by the highway itself (geometric elements, road surface, ancillary facilities, etc.). The limited speed refers to the maximum operating speed allowed on the highway; The desired speed refers to the speed that the vehicle can maintain without being restricted by other surrounding vehicles. In this paper, small car is selected as evaluation model. According to the SHSA, the desired speed is shown in Table 2.

3 Design Consistency Evaluation Model

3.1 Speed Consistency Evaluation Model in SHSA

The SHSA gives a speed consistency evaluation model based on operating speed, prescribed by the standard of the highway road sections and different speed prediction model, and provides the speed consistency by the absolute value of adjacent sections

Table 3 Highway section division standard

Alignment	Horizontal curve radius: $R > 1000$ m	Horizontal curve radius: $R \leq 1000$ m
Vertical grade: $G < 3\%$	Tangent length: $L > 200$ m Horizontal tangent section	Horizontal curve section
	Tangent length: $L \leq 200$ m Horizontal tangent section	
Vertical grade: $G \geq 3\%$	Vertical slope section	Horizontal curve-vertical slope combined section

Table 4 Speed consistency evaluation criteria in SHSA

Speed consistency	$ \Delta V_{85} $	$ \Delta V $
Good	$ \Delta V_{85} < 10$ km/h	$ \Delta V < 10$ km/h
Fair	$10 \text{ km/h} \leq \Delta V_{85} < 20$ km/h	$10 \text{ km/h} \leq \Delta V < 20$ km/h
Poor	$ \Delta V_{85} \geq 20$ km/h	$ \Delta V \geq 20$ km/h

speed difference $|\Delta V_{85}|$ and the absolute value of the same section of design speed and operating speed difference $|\Delta V|$, road division standard are shown in Table 3, the speed consistency evaluation standard are shown in Table 4.

3.2 Design Consistency Evaluation Model in IHSDM

In IHSDM, the consistency of the design is evaluated by establishing the operating speed analysis model along the highway, and the model is first combined with the horizontal and vertical alignment to predict the operating speed of the tangent section, the horizontal curve section, the horizontal curve and the vertical curve combined section (see Table 5). Then, predicts the rates of acceleration and deceleration in and out of the horizontal curve by the prediction formula (see Table 6). Finally, adjusts the operating speed of the tangent section by acceleration and deceleration rates of the section in and out of the horizontal curve, and generate the final operating speed (V_{85}) profile.

In the horizontal curve section and the combination of horizontal curve and vertical curve section, the operating speed prediction formula is mainly related to the radius R of the horizontal curve. Only when the longitudinal slope grade is different, the formula will be different, but the formula structure remains the same. However, under actual conditions, the speed of vehicles on the expressway is not only determined by the radius of the horizontal curve, but also by the horizontal and vertical alignment. The design consistency evaluation model ignores the influence of the vertical slope gradient below 4% on the operating speed and the influence of the vertical curve radius on the operating speed. Therefore, the design consistency evaluation model

Table 5 Operating speed prediction formula

Equation	Alignment type/condition	V_{85} (km/h)
1	Tangent	$V_{85} = V_{\text{desire}}$
2	Horizontal curve on grade: $G < -4\%$	$V_{85} = 102.10 - 3077.13/R$
3	Horizontal curve on grade: $-4\% \leq G < 0\%$	$V_{85} = 105.98 - 3709.90/R$
4	Horizontal curve on grade: $0\% \leq G < 4\%$	$V_{85} = 104.82 - 3574.51/R$
5	Horizontal curve on grade: $G \geq 4\%$	$V_{85} = 96.61 - 2752.19/R$
6A	Horizon curve and vertical curve combined: vertical curve begins before midpoint of horizontal curve	Calculate “effective grade” and use appropriate Eqs. 2–5
6B	Horizon curve and vertical curve combined: vertical curve begins after midpoint of horizontal curve	Use Eqs. 2–5 based on entry grade

Table 6 Acceleration and deceleration rates prediction formula

Classify	Condition	Horizontal curve radius (R) (m)	Rates (m/s ²)
Deceleration	Tangent to curve	$R > 873$	0.05
		$175 \leq R \leq 873$	$-0.0008726 + 37,430/R^2$
		$R < 175$	1.25
	Curve to tangent, or tangent to tangent	All	0.05
	To a STOP condition		2.5
	To an end speed (>0 km/h)		1.25
Acceleration	Curve to tangent	$R > 436$	0.21
		$250 \leq R \leq 436$	0.43
		$R < 250$	0.54
	Tangent to curve, or tangent to tangent	All	0.21
	From a STOP condition		1.54
	From a start speed (>0 km/h)		0.54

in IHSDM is not easy to find the sections with problems in the vertical alignment, resulting in poor design safety evaluation effect of the vertical alignment.

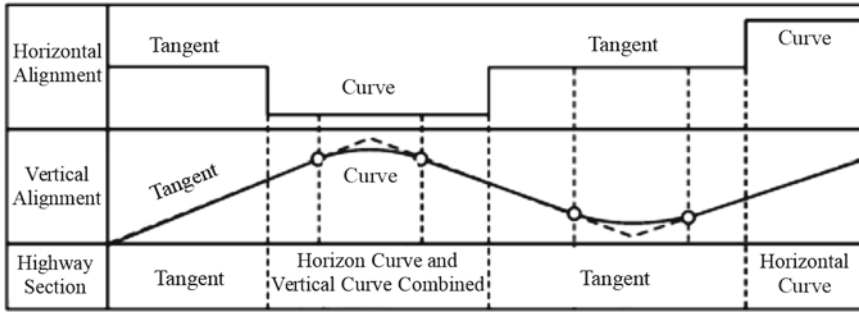


Fig. 1 Schematic diagram of road section division

3.3 Improved Design Consistency Evaluation Model

Based on the actual highway design data in China, improved the design consistency evaluation model in IHSDM, The operational steps to improve the model are as follows:

First of all, according to the horizontal and vertical alignment of the highway, the section is divided into tangent section, horizontal curve section, the horizontal curve and the vertical curve combined section (see Fig. 1).

Secondly, the operating speed of each section is predicted according to the improved speed prediction formula. It is assumed that the operating speed of the tangent segment can reach the desired speed.

In combination with SHSA, the operating speed prediction formula of the horizontal curve section and the horizontal curve and the vertical curve combination section is improved, and the improved prediction formula is as follows:

$$V_1 = \begin{cases} 102.10 - 3077.13/R, & G \leq -4\%. \\ 103.23 - 3168.35/R, & -4\% \leq G < -3\%. \\ 105.98 - 3709.90/R, & -3\% \leq G < 0\%. \\ 104.82 - 3574.51/R, & 0\% \leq G < 3\%. \\ 99.33 - 2904.85/R, & 3\% \leq G < 4\%. \\ 96.61 - 2752.19/R, & G \geq 4\%. \end{cases} \quad (1)$$

- V_1 the operating speed on the horizontal curve (km/h);
- R the horizontal curve radius (m);
- G the longitudinal slope grade (%).

According to the radius of the vertical curve, the operating speed of the combined section of the vertical curve and the flat curve is adjusted, and the formula for calculating the adjustment value is shown in Table 7.

The effect of the vertical curve on the operating speed is mainly related to the radius of the vertical curve. The smaller the radius of the vertical curve is, the lower

Table 7 Calculation formula of operating speed adjustment value

V_{design} (km/h)	ΔV (km/h)	
	Concave vertical curve	Convex vertical curve
120	$\Delta V = \frac{17,000-R_1}{600}$	$\Delta V = \frac{6000-R_1}{200}$
100	$\Delta V = \frac{10,000-R_1}{350}$	$\Delta V = \frac{4500-R_1}{150}$
80	$\Delta V = \frac{4500-R_1}{150}$	$\Delta V = \frac{300-R_1}{100}$

the operating speed, and vice versa. When the vertical curve radius is relatively large and reaches a certain value, the influence on the running speed can be ignored, and the adjustment formula is as follows:

$$V'_{85} = \begin{cases} V_{85} - \Delta V, & \Delta V \geq 0 \\ V_1, & \Delta V < 0 \end{cases} \tag{2}$$

R_1 the vertical curve radius (m);

V'_{85} the adjusted speed in horizontal curve and vertical curve combined section (km/h).

Thirdly, predict the rates of acceleration and deceleration in and out of the horizontal curve by the prediction formula in IHSDM. and adjust the operating speed of the tangent section by acceleration and deceleration rates of the section in and out of the horizontal curve to obtain the final operating speed curve.

According to Formulas 3 and 4, the required distances for acceleration and deceleration are calculated and compared with the tangent length between curve segments.

$$S_n = \frac{|V_{\text{desire}}^2 - V_n^2|}{2 \times a_n \times 3.6^2} \tag{3}$$

$$S_{n+1} = \frac{|V_{\text{desire}}^2 - V_{n+1}^2|}{2 \times a_{n+1} \times 3.6^2} \tag{4}$$

S_n the required distances for acceleration (m);

S_{n+1} the required distances for deceleration (m);

a_n acceleration rates from curve n to tangent (m/s²);

a_{n+1} deceleration rates from tangent to curve $n + 1$ (m/s²);

V_n operating speed on the curve n (km/h);

V_{n+1} operating speed on the curve $n + 1$ (km/h).

When the tangent length between the curve segments is not less than the sum of the acceleration and deceleration required distances, the vehicle can accelerate to the desired vehicle speed on the straight segment, the operating speed adjustment curve is shown in Fig. 3b; and when the tangent length between the curve segments is less than the sum of the acceleration and deceleration required distances, the desired

vehicle speed cannot be accelerated and the operating speed adjustment curve is shown in Fig. 3c.

The formula for calculating the operating speed of the starting deceleration point is as follows:

$$V = 3.6 \times \sqrt{\frac{2a_n \times a_{n+1} \times l + a_{n+1} \times \frac{V_n^2}{3.6^2} + a_n \times \frac{V_{n+1}^2}{3.6^2}}{a_n + a_{n+1}}} \tag{5}$$

V the operating speed of the starting deceleration point (km/h);

l the length of tangent between curve n and curve $n + 1$ (m) (Fig. 2).

Finally, according to the obtained curve graph of the final operating speed, the design consistency evaluation index is used to evaluate the running safety of the expressway. There are two ways to evaluate the design consistency: the difference between the operating speed and the design speed of the same section, and the difference between the operating speed of different adjacent section (horizontal tangent and curve). According to the above two evaluation indicators, judge whether the highway design is continuous, so as to evaluate the road safety. Design consistency evaluation criteria are expressed by indicators ΔP_1 (km/h) and ΔP_2 (km/h), as is shown in Table 8.

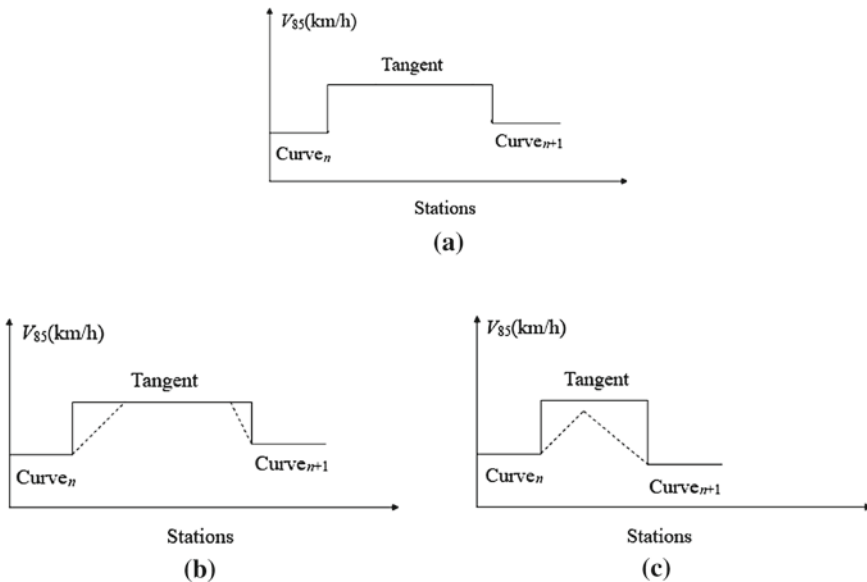


Fig. 2 Operating speed adjustment diagram

Table 8 Evaluation indicators for design consistency

Design consistency	$\Delta P_1 = V_{85} - V_{design}$	$\Delta P_2 = V_{85Tangent} - V_{85curve}$
Good	$0 \leq \Delta P_1 \leq 10$	$\Delta P_2 \leq 10$
Fair	$10 < \Delta P_1 \leq 20$	$10 < \Delta P_2 \leq 20$
Poor	$20 < \Delta P_1$	$20 < \Delta P_2$
Good	$\Delta P_1 < 0$	–

4 Case Analysis

The example adopted in this paper is a two-way four-lane expressway with a design speed of 120 km/h in China. The improved model was used to evaluate the design consistency, according to Table 1, it can be determined that the desired speed of this expressway is 120 km/h. The expressway is divided into 16 horizontal tangent and curve section. The operating speed of each section was predicted and adjusted to obtain the final operating speed curve. Finally, the safety evaluation of the expressway was conducted according to the design consistency evaluation index. The operating speed curve is shown in Fig. 3, and the design consistency evaluation results are shown in Table 8.

The Design Specification for Highway Alignment (DSHA) in China is adopted to evaluate the alignment design of the expressway, and the evaluation results of the DSHA are compared with the design consistency evaluation results to verify the effectiveness of the design consistency evaluation results. The evaluation results of the DSHA showed that there were 6 sections of the highway with deficiencies in alignment design, among which 5 sections were evaluated as “fair” or “poor” by the improved model, and the accuracy of the evaluation results was 83.3%, the improved model can effectively evaluate the safety of expressway operation (Table 9).

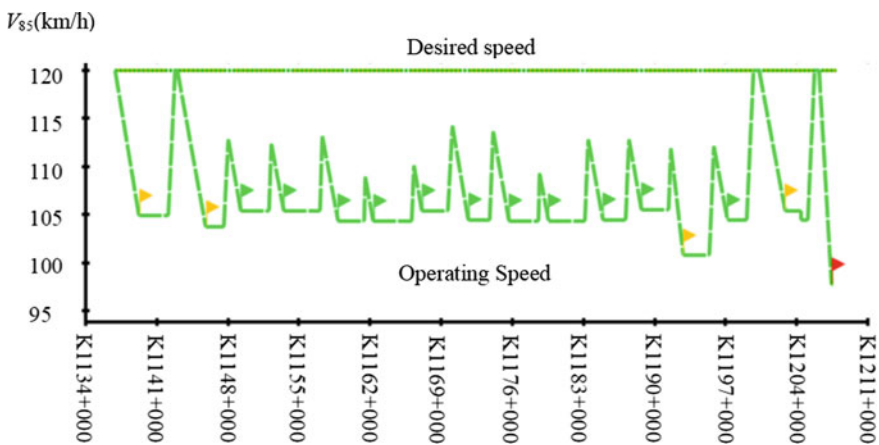


Fig. 3 Operating speed profile

Table 9 Design consistency evaluation results

No	Horizontal alignment	Pile number	V_{85} (km/h)	ΔP_1 (km/h)	ΔP_2 (km/h)	Design consistency
1	Tangent	K1137 + 000 ~ K1139 + 331	120	0	15	Fair
	Curve	K1139 + 331 ~ K1142 + 187	105	-15		
2	Tangent	K1142 + 187 ~ K1145 + 915	120	0	16	Fair
	Curve	K1145 + 915 ~ K1147 + 690	104	-16		
13	Tangent	K1173 + 786 ~ K1175 + 721	111	-9	10	Fair
	Curve	K1175 + 721 ~ K1178 + 574	101	-16		
15	Tangent	K1199 + 109 ~ K1202 + 866	120	0	15	Fair
	Curve	K1202 + 866 ~ K1205 + 205	105	-15		
16	Tangent	K1205 + 205 ~ K1207 + 492	120	0	22	Poor
	Curve	K1207 + 492 ~ K1207 + 646	98	-22		

5 Conclusion

Based on the actual highway design data in China, this paper improves the design consistency evaluation model in IHSDM, and proposes an alignment operation safety evaluation method for expressways in consideration of design consistency, determines the indexes of highway safety evaluation and the method of section division. Combined with the SHSA and IHSDM, a model of highway speed prediction and design consistency evaluation indicators suitable for China was constructed, an example is given to verify the effectiveness of the model. Study on the evaluation of highway alignment operation safety considering design consistency can provide basis for quantifying the security of highway alignment design during the design, construction and operation stages.

Acknowledgements This research was supported by National Key Research and Development Program of China (2017YFC0803901), Heilongjiang highway survey and design institute project (2018006).

References

1. Ma C, Zhang S, Ma Z (2018) Nonlinear negative binomial regression model of expressway traffic accident frequency prediction. *China J Highw Transp* 31(11):180–189
2. Hassan Y, Gibreel G, Easa SM (2000) Evaluation of highway consistency and safety: practical application. *J Transp Eng* 126(3):193–201
3. Castro M, Sanchez JF, Sanchez JA et al (2011) Operating speed and speed differential for highway design consistency. *J Transp Eng* 137(11):837–840
4. Liu Z, Chen J, Song S (2008) Development and application of two-lane highway design consistency evaluation software. *J Highw Transp Res Dev* 6:47–51
5. Camacho-Torregrosa FJ, Pérez-Zuriaga AM, Campoy-Ungría JM et al (2012) New geometric design consistency model based on operating speed profiles for road safety evaluation. *Accid Anal Prev* 61(8):33–42
6. Gao J, Guo Z (2004) Evaluation of highway linear design quality based on running speed. *J Tongji Univ* 32(7):906–911
7. Liu J, Guo Z, Hu J, Zhang Y (2010) Safety evaluation method and standard of highway alignment design. *J Highw Transp Res Dev* 23(S1):28–35
8. Yang H, Zhang J, Xu J (2010) Horizontal alignment safety evaluation of two-lane highway based on operating speed. *J Highw Transp Res Dev* 27(09):127–131+142
9. Wang S, Guo T, Liu J, Luo M (2010) Research on China highway operating speed system and engineering application technology. *China J Highw Transp* 23(S2):1–7+23
10. Deng Y (2009) Speed prediction model of free-flowing passenger car on long and steep downgrades. *J Chang'an Univ* 29(04):43–47
11. He Y, Wang S, Sun X, Zhang J, Hou S (2010) Operating speed and highway safety in China and the United States. *China J Highw Transp* 23(S1):73–78
12. Wang S, Fang J, Zhou R, Zhu Z (2010) Research on operating speed characteristics of highway. *China J Highw Transp* 23(S1):24–27+46
13. Guo Q, Wang X, Chen Z (2019) Modeling operation speed on mountainous freeways: a driving simulator study. *J Tongji Univ* 47(07):1004–1010
14. Huang J (2003) Introduction to the interactive highway safety design model in United States. *J Highw Transp Res Dev* 05:115–117
15. Meng X, Hou Q, Shi Y, Jing L, Zheng L (2016) IHSDM freeway accident prediction models. *J Traffic Transp Eng* 16(01):123–132
16. Duan M, Tang B, Liu F (2008) Study on the applicability of IHSDM in accident prediction of freeway of high proportion of bridges and tunnels in mountain areas. *Highw Eng* 43(03):70–76
17. Liu Z, Wang Y, Huang C (2008) Application of interactive road safety design model IHSDM and design consistency model DCM. *J Highw Transp Res Dev* 06:42–45
18. Fu Z, Liu Z (2010) Prediction model of theoretical operating speed based on combination alignment of cross and vertical section. *J Chang'an Univ* 30(03):25–27
19. Zhou R, Sun J, Wu W (2003) Relationship between running speed and gradient of expressway. *J Highw Transp Res Dev* 20(4):34–37

Forecast Model of Comprehensive Transport Freight Index Considering Macro Economy in China



Yanli Ma, Jieyu Zhu, and Yining Lou

Abstract To better construct the coordinated development of the comprehensive transportation system, the prediction model of comprehensive freight index in China was studied. Basic indicators of production and supply, benefits and macroeconomic were selected. Cluster analysis was used to divide the indicators into 13 leading indicators and 13 synchronous indicators. Factor analysis was used to analyze the importance of each leading indicator. The characteristic quantities of freight index were determined and the Granger causality test was conducted between the four indexes and the five characteristic quantities respectively. The relationship model between each freight index and characteristic quantity was constructed, and the validity of the model was verified. Results show that the freight indexes are volatile. There is a stable relationship between characteristic quantity and freight index. The error of forecast and actual value of is within 10%, verifying the validity of the model. The results can provide theoretical support for forecasting the development trends of freight transport among various transport modes. The comprehensive transport will be improved to provide a decision-making basis for the national macro-department to make policies.

Keywords Comprehensive transportation · Freight index · Clustering analysis · Granger causality test · Forecast model

1 Introduction

As the basic industry of national economy and social development, the transportation industry is one of the important indexes to judge the trend of national economy development and reflect the periodicity of economy. It should give full play to its role

Y. Ma (✉) · J. Zhu

School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China

e-mail: mayanli@hit.edu.cn

Y. Lou

Department of Mathematics, University College London, London WC1E 6BT, UK

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

231

W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,

Lecture Notes in Electrical Engineering 775,

https://doi.org/10.1007/978-981-16-5429-9_16

in leading economic and social development. As an important part of transportation, comprehensive cargo transportation is closely related to economic development. At present, China's economy is in the stage of new normal development, and the freight structure of various modes is constantly changing. Under the development of internet platforms, the demand for transportation also keeps growing. To better meet the requirements of the balance of supply and demand, and the healthy development of the national economy, it is urgent to establish an efficient transportation system integrating transportation services. Therefore, it is particularly important to study the forecast model of China's comprehensive transport freight index to reveal the development trend of China's comprehensive transport.

To accurately analyze operation of freight transportation system, several studies discussed the impact of economic growth on transportation. Kuzu and Önder [1] investigated the long-term relationship between the development of the logistics industry and economic growth. GDP data were used to measure economic growth, whereas the transport warehouse turnover was an indicator of logistic development. Results showed that economic growth plays an active role in logistics development. Gao et al. [2] analyzed the intrinsic relationship between comprehensive transportation freighting index and GDP to detect the problems in the development of transportation. Results showed that the volume of freight traffic and freight turnover is positively correlated with GDP. Guo and Xia [3] analyzed the relationship between traffic facilities and economic growth based on the general VAR model using the GDP data of China, railway and highway mileage from 1964 to 2004. Results showed that the selected variables are co-integrated with positive and negative contributions of railway and highway investments to GDP, respectively. Muller et al. [4] examined the impact of air, railway and container port transportation on economic growth. Results indicated that container port traffic positively influenced per capita income across countries. Meng et al. [5] adopted statistical methods to explore the internal connection between regional railway freight volume and regional GDP. They showed that the relationship between railway transportation and economic growth was mutually promoting. Pandawa et al. [6] studied the relationship between freight and economic growth adopting three-stage least square method for estimation. Results showed a bidirectional causality relationship between economic growth and freight transport in high- and low-income countries. Pradhan [7] examined the interactions between transportation infrastructure and economic growth in the G-20 countries. Results showed that economic growth causes transportation infrastructure. Lei [8] established a coupling model to study the interaction between the traffic and regional economy of the Hui community. Tong and Yu [9] investigated relationship between freight and economic growth in the underdeveloped central and western regions. Empirical results showed that the interaction between economic growth and transportation services was generally positive. Jiang et al. [10] proposed a structural equation model to consider the relationship between multimodal transportation and economic development comprehensively. Travel demand was added as an endogenous variable to account for the complicated interactions between transportation and economic.

Some scholars have studied freight indexes. Okan and Yoshida [11] proposed to establish a long-term dry bulk freight index and studied its characteristics in periodic fluctuations. They used the data to build a model of long-term freight rate and maritime trade, and took life expectancy as a long-term leading indicator. Kumar and Anbanandam [12] proposed a social sustainability index calculation framework based on the dimensions and attributes of social sustainability promoting factors in goods transportation. The framework considered 16 dimensions and 74 attributes of 4 factors, including 17 new attributes of social sustainability, and calculated the overall social sustainability index based on fuzzy logic. Tsioumas et al. [13] proposed a VARX model based on exogenous variables, which combined the development of China's steel production dry bulk fleet and the new comprehensive dry bulk Economic climate index (DBECI). The results showed that the selected independent variables could improve the accuracy of BDI prediction. Duru et al. [14] proposed a bivariate long-term fuzzy inference system for time series forecasting task in the field of freight market. Empirical study of shipping rate data was investigated and takes life expectancy was considered as the leading index of time series model. Angelopoulos et al. [15] used monthly composite index and proprietary annual Freight rate Index (LFI) to investigate the long-term and short-term dynamic spectrum dry cargo shipping market, and evaluated signs of long-term cyclical effects in the shipping industry through LFI. Khan and Khan [16] provided an empirical analysis of rail freight transport demand in Pakistan using annual time series data from 1972 to 2017. The Johansen's co-integration and error-correction model was used to estimate short and long-run elasticities. Moscoso-López et al. [17] presented and compared two forecasting-models to predict the freight volume. The models developed and tested are based on Artificial Neural Networks and Support Vector Machines. Uyar et al. [18] presented a genetic algorithm based trained recurrent fuzzy neural network for long term dry cargo freight rates forecasting. Betanzo et al. [19] estimated a hedonic pricing regression to generate a market index from heterogeneous fixture data in the Offshore Support Vessel (OSV) market. Results show that the volatile market index was seasonal and was positively correlated to both oil prices and production volumes. Betanzo and Romero [20] presented a multi-factorial model to assess the performance of authorities in dealing with the issues derived from freight transport within cities. The model produces an Urban Freight Transport Index (UFTI), which quantifies the intention of the authorities to cope with urban freight transport externalities. Liu et al. [25] designed a road freight index system based on the current situation of Ningbo road freight transport market for setting the freight index in Ningbo. Models to calculate the average freight indicator are formulated from different perspectives. Løvold Rødseth [21] established a logistics productivity index that decomposes into technical, cargo mix, vehicle capacity, and efficiency changes to determine the rate and drivers of growth. Gong et al. [22] built a railway coal transport boom indicator system with 108 monitoring indicators based on the analysis of the supply and demand of the coal market. The diffusion index and synthesis index are calculated, judging the trend of railway coal transport market comprehensively. Shi and Yang [23] put forward that measuring the development level of a single mode of transport generally used its own vertical prosperity index. Ning et al. [24] used factor analysis

to compile global shipping boom diffusion index and composite index, evaluating the dynamic change process of global shipping boom.

The above scholars mainly focus on the research of relevant indicators of different modes of transport, and mainly in the aspect of freight price index. There are relatively few researches on the comprehensive transport freight index, so the forecast model of the comprehensive transport freight index was studied. The freight index in this paper is the ratio of the amount of goods shipped by various modes of transport to the total amount of shipped goods. It reflects the status of the transportation mode in the integrated transportation system. The importance of each leading indicator was analyzed through basic indicators selection and factor analysis method. Then the relationship model between each freight index and characteristic quantity was constructed. The research results can provide a theoretical basis for the forecast and development trends of freight transportation, which can accelerate the construction of an integrated transportation network.

2 Indicator Selection and Characteristic Quantity Determination

2.1 Construction of Freight Indicator System

Choosing the right indicator is the key of forecast model the comprehensive freight index. Its integrity and non-repeatability of reflecting the market activity should be considered. According to the principle of systematization, synchronization, comparability, sensitivity and smoothness of operability, the 26 basic indicators were selected from 4 categories (production category, supply category and benefit category, macroeconomic category), as shown in Table 1.

2.2 Division of Freight Indicators

To meet the demand of freight index forecast, the 26 indicators were reclassified by cluster analysis into leading and synchronous indicator. Cluster analysis can automatically classify the sample data, mainly including splitting method, hierarchical method, density-based method and grid-based method. Rapid clustering and hierarchical clustering analysis were adopted to divide the object set with the structure of tree. Finally, other clustering methods were used to optimize these clusters and divide the indicator types. To avoid strong collinearity between variables, the advance preprocessing of the cluster analysis was conducted. There was a multi-collinearity relationship through correlation analysis of 26 indicators. The car ownership indicators of national railway freight were eliminated and the remaining indicators were

Table 1 Freight indicator system of comprehensive transport

Indicator system	Railway	Highway	Water transport	Aviation
Production	Freight volume (X_1)	Freight volume (X_2)	Freight volume (X_3)	Freight volume (X_4)
	Freight turnover (X_5)	Freight turnover (X_6)	Freight turnover (X_7)	Freight turnover (X_8)
Supplies	Truck ownership (X_9), locomotive ownership (X_{10})	Truck ownership (X_{11})	New capacity in water transport (X_{12})	Motor vehicle ownership (X_{13})
	Revenue kilometers (X_{14})	Highway mileage (X_{15})	Mileage of inland waterways (X_{16})	Total length of air mail route (X_{17})
Benefits	Fixed investments (X_{18})	Fixed investments (X_{19})	Fixed investments (X_{20})	Fixed investments (X_{21})
Macro economy	Total amount of import and export (X_{22})	Added value of first industry (X_{23})	Added value of secondary industry (X_{24}), added value of the tertiary industry (X_{25})	Consumer price index (X_{26})

classified into the category of highway freight volume. The cluster analysis tree is shown in Fig. 1.

The indicators were divided into two categories through cluster analysis of software SPSS. The clustering results were shown in Table 2.

2.3 Weight Analysis Freight Indicator

Principal component analysis (PCA) was used to extract common factor from the leading indicator data set, and orthogonal rotation was taken as factor rotation. The most commonly used theoretical model to explain the total factor analysis was obtained as follows.

$$X_j = a_{j1}F_1 + a_{j2}F_2 + \dots + a_{jm}F_m + U_j \tag{1}$$

where F is the common factor of each indicator component, the mean value and variance of F_m were 0 and 1, and they are independent of each other; $j = 1, 2, 3, \dots, n$, n is the number of indicators; is a specific factor, it only works for X_j , whose mean is 0. Each U_j is independent of each other, which is equivalent to the residual in multiple linear regression model. The weight determination method of index X_j was as follow.

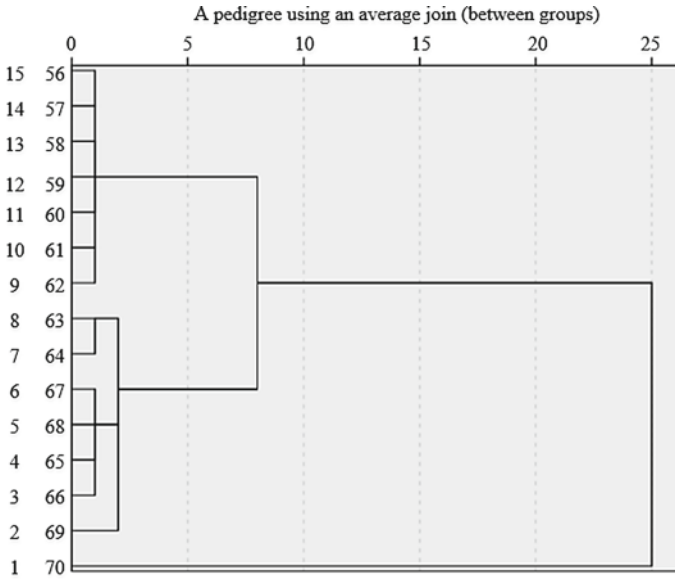


Fig. 1 Cluster analysis tree

Table 2 Cluster analysis results

Indicator type	Indicator notation
Leading indicators	$X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{13}, X_{15}, X_{16}, X_{17}, X_{18}, X_{21}$
Synchronized indicators	$X_1, X_2, X_3, X_4, X_{12}, X_{14}, X_{19}, X_{20}, X_{22}, X_{23}, X_{24}, X_{25}, X_{26}$

$$w_j = \frac{\sum_{i=1}^m a_{ji}^2}{\sum_{j=1}^n \sum_{i=1}^m a_{ji}^2} \tag{2}$$

where m is the number of factors, w_j is the weight of each indicator.

To determine whether the leading indicator was suitable, KMO and Bartlett tests were performed, and the test results were shown in Table 3.

As depicted in Table 3, KMO information was 0.797, KMO was the sampling suitability quantity of Kaiser–Meyer–Olkin. When the value of KMO measure was

Table 3 KMO and Bartlett tests of leading indicators

<i>Kaiser–Meyer–Olkin</i> measurement with sufficient sampling		0.797
<i>Bartlett’s</i> sphericity test	Approximate chi-square	290.821
	<i>Df</i>	78
	<i>Sig</i>	0.000

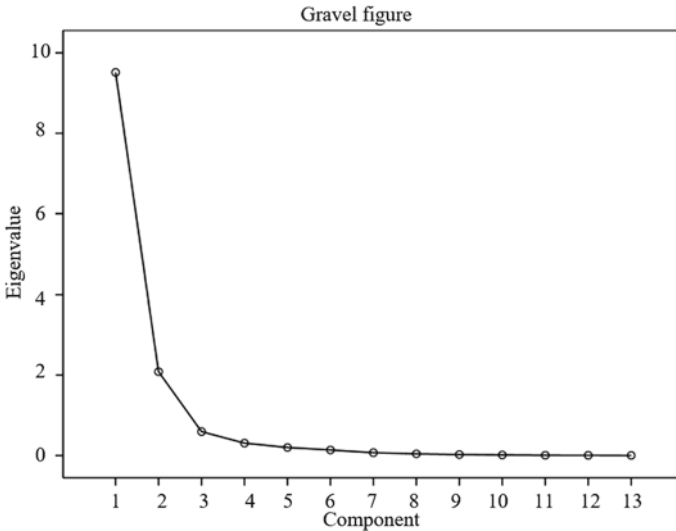


Fig. 2 Lithotripsy diagram of leading indicator factor analysis

greater than 0.6, it indicates that the indicator data was suitable for factor analysis. The approximate chi-square of Bartlett sphericity test was 290.821, the degree of freedom was 78, and the significance probability of the test was 0.000, indicating that the data was suitable for factor analysis.

Through factor analysis of the leading indicators, the gravel diagram of factor analysis was shown in Fig. 2. The lithotripsy diagram of the first two common factors was relatively steep and then tends to be flat, so they reflected most of the information.

The weight of leading indicators can be obtained by using formula 2, and the results were shown in Table 4.

From Table 4, railway freight turnover (X_5), highway freight turnover (X_6), highway truck ownership (X_{11}), civil motor vehicle ownership (X_{13}) and inland waterway mileage (X_{16}) were selected as the characteristic quantities of the freight indicators.

3 Forecast Model of Freight Index

3.1 Data Processing

Data of various indicators from 2000 to 2017 were obtained from the National Bureau of Statistics. X_5 stands for rail freight turnover (10^8 t km), X_6 is highway freight turnover (10^8 t km), X_{11} stands for highway truck ownership (ten thousand car),

Table 4 Weight results of leading indicators

Indicator	Component		Weight
	1	2	
Cargo turnover of civil aviation (X_8)	0.986	0.050	7.559898
Investment of fixed assets in the air transport industry (X_{21})	0.965	-0.225	6.938573
Water freight turnover (X_7)	0.959	0.099	2.849937
Highway freight turnover (X_6)	0.947	0.233	12.04686
Highway truck ownership (X_{11})	0.934	0.257	14.43341
Investment of fixed assets in railway transportation (X_{18})	0.923	0.261	7.897164
Total length of air mail route (X_{17})	0.919	-0.192	4.757622
Mileage of inland waterways (X_{16})	0.902	-0.149	9.715887
highway mileage (X_{15})	0.892	0.253	6.768905
Ownership of civilian motor (X_{13})	-0.829	0.315	12.11697
Wagons ownership of national railway (X_9)	-0.370	0.797	3.722584
Locomotive ownership of national railway (X_{10})	-0.591	0.728	3.259887
Freight turnover of the railway (X_5)	0.666	0.667	7.932307

X_{13} stands for civilian motor vehicle ownership (ship), X_{16} is the mileage of inland waterways (10^4 km). RX_1 represents the freight index of railway, RX_2 represents the freight index of highway, RX_3 represents the freight index of water transport, RX_4 represents the freight index of aviation. The trends of each indicator and freight index were shown in Figs. 3 and 4.

3.2 Granger Causality and Stationary Test

To explain whether each characteristic quantity can explain each transport mode's freight index, Granger causality test was conducted between the four freight indexes and the five characteristic quantities. Granger causality test results of railway freight index and each characteristic quantity were shown in Table 5. Similarly, Granger causality test results of highway, waterway freight index, air freight index and each characteristic quantity can be obtained. Due to limited space, their tables were omitted here.

According to the Granger test, Granger causality exists in X_{10} and RX_1 as well as X_{15} and RX_1 in terms of railway freight index, so VAR model can be established. Granger causality exists between X_{10} and RX_2 , X_{14} and RX_2 , X_{15} and RX_2 and X_{25} and RX_2 of highway freight index. Granger causality exists among X_{10} and RX_2 , X_{14} and RX_2 , X_{15} and RX_2 as well as X_{25} and RX_2 of water freight index, so VAR model can be considered. In terms of air freight index, there was no Granger causality between RX_4 and other characteristic quantities, but it may be caused by little data,

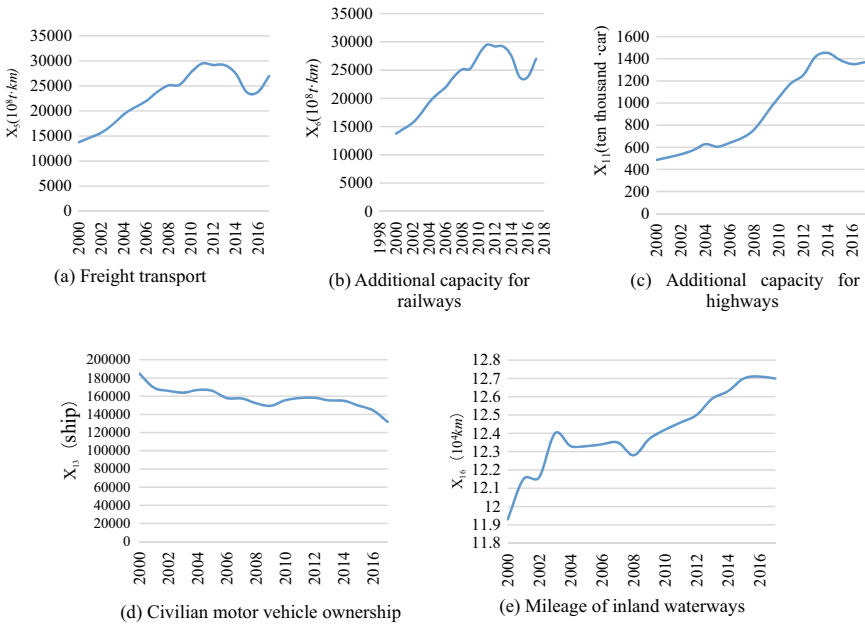


Fig. 3 Changes of each characteristic quantity from 2000 to 2017

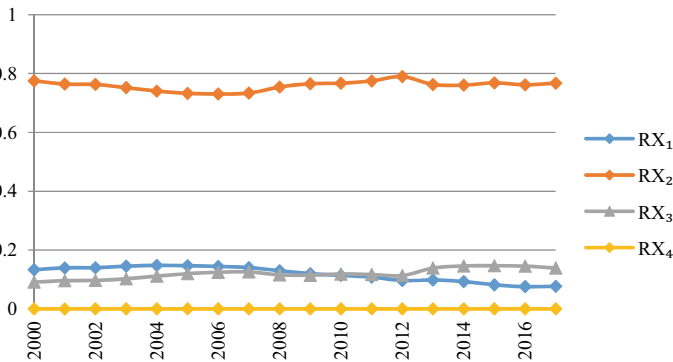


Fig. 4 Changes of freight index from 2000 to 2017

which was not a necessary condition for establishing VAR model, so it can continue to model.

LagLengthCriteria was adopted to determine the lag period, and the maximum lag period of freight index of railway, highway, water transportation and air transportation were 1. By testing the stability of freight index and characteristic quantity of each mode of transport, it was found that the railway freight index, road freight index, water freight index and air freight index and characteristic quantity sequence meet the stable

Table 5 Granger causality test results of railway freight index and each characteristic quantity

Dependent variable	Excluded	Chi-sq	df	Prob.
X_5	RX_1	1.776119	1	0.1826
RX_1	X_5	1.151215	1	0.2833
X_6	RX_1	5.745235	1	0.0165
RX_1	X_6	1.273201	1	0.2592
X_{11}	RX_1	0.747181	1	0.3874
RX_1	X_{11}	0.606542	1	0.4361
X_{13}	RX_1	3.233244	1	0.0722
RX_1	X_{13}	2.744000	1	0.0976
X_{16}	RX_1	0.482934	1	0.4871
RX_1	X_{16}	0.697758	1	0.4035

conditions of VAR model. A vector auto-regression model can be established. The freight index of each transport mode and the stationarity results of each characteristic quantity was shown in Fig. 5.

3.3 Establishment of VAR Model

GDP and freight volume model orders were determined through test results. The VAR model parameter of railway freight index were shown in Table 5. The results include three parts: parameter estimation of the model, standard deviation of estimated coefficient and t-test statistical value. Similarly, the VAR model parameter results of highway, waterway and air freight indexes can be determined.

The freight index calculation model can be obtained through coefficient estimation results. The freight index model of each transportation mode was shown in formula (3)–(6).

$$\begin{aligned}
 RX_1 = & -1.18E - 06 \times X_{5(t-1)} + 3.93E - 07 \times X_{6(t-1)} \\
 & - 6.17E - 07 \times X_{11(t-1)} + 3.12E - 07 \times X_{13(t-1)} \\
 & - 1.52E - 08 \times X_{16(t-1)} + 1.21 \times RX_{1(t-1)} - 0.061113 \quad (3)
 \end{aligned}$$

$$\begin{aligned}
 RX_2 = & 1.01E - 06 \times X_{5(t-1)} - 6.14E - 07 \times X_{6(t-1)} \\
 & + 6.17E - 05 \times X_{11(t-1)} - 9.86E - 07 \times X_{13(t-1)} \\
 & - 0.09 \times X_{16(t-1)} + 0.62 \times RX_{2(t-1)} + 1.503920 \quad (4)
 \end{aligned}$$

$$\begin{aligned}
 RX_3 = & 1.05E - 07 \times X_{5(t-1)} + 3.45E - 07 \times X_{6(t-1)} \\
 & - 2.10E - 05 \times X_{11(t-1)} + 6.25E - 07 \times X_{13(t-1)}
 \end{aligned}$$

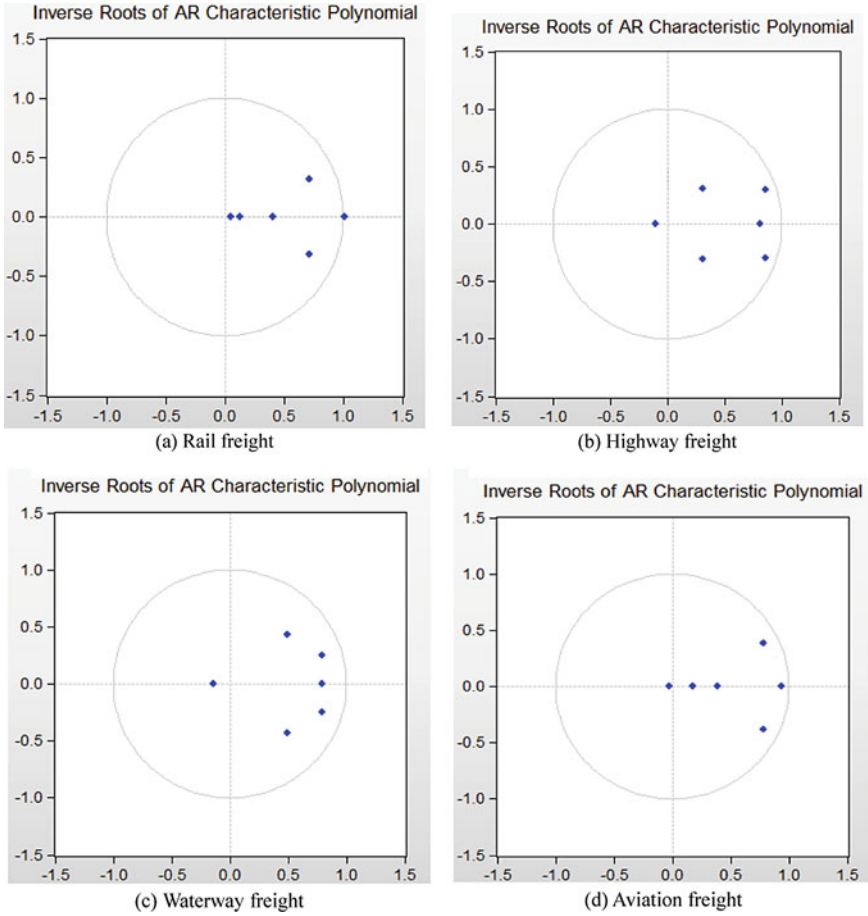


Fig. 5 Results of exponential stationarity test for each mode of transport

$$+ 0.05 \times X_{16(t-1)} + 0.61 \times RX_{3(t-1)} - 0.703544 \tag{5}$$

$$\begin{aligned}
 RX_4 = & 1.59E - 10 \times X_{5(t-1)} - 3.39E - 10 \times X_{6(t-1)} \\
 & - 3.21E - 08 \times X_{11(t-1)} + 8.61E - 10 \times X_{13(t-1)} \\
 & + 6.44E - 05 \times X_{16(t-1)} + 0.29 \times RX_{4(t-1)} - 0.000514 \tag{6}
 \end{aligned}$$

where RX_1 is Railway freight index, $X_{5(t-1)}$ is rail freight turnover for $(t - 1)$ -th year (10^8 t km), $X_{6(t-1)}$ is highway freight turnover for $(t - 1)$ -th year (10^8 t km), $X_{11(t-1)}$ is highway truck ownership for $(t - 1)$ -th year (ten thousand car), $X_{13(t-1)}$ is Civilian motor ownership for $(t - 1)$ -th year (ship), $X_{16(t-1)}$ is mileage of inland waterways (10^4 km) for $(t - 1)$ -th year, $RX_{1(t-1)}$ is rail freight index for $(t - 1)$ -th year, RX_2 is road freight index, $RX_{2(t-1)}$ is road freight index for $t - 1$ year, RX_3 is

Table 6 Estimated model parameters of railway freight index

	X5	X6	X11	X13	X16	RX1
X5 (-1)	1.145689	2.898098	0.014607	-1.127537	3.492598	-1.18E-06
	(0.22148)	(0.88039)	(0.00664)	(0.72290)	(4.60889)	(8.4E-07)
	[5.17283]	[3.29182]	[2.20069]	[-1.55975]	[0.75780]	[-1.40209]
X6 (-1)	-0.037249	-0.291852	0.005150	0.599885	-4.600618	3.93E-07
	(0.10956)	(0.43551)	(0.00328)	(0.35760)	(2.27990)	(4.2E-07)
	[-0.33998]	[-0.67014]	[1.56861]	[1.67754]	[-2.01790]	[0.94547]
X11 (-1)	-7.990672	-25.04366	0.365709	16.08432	185.0295	-6.17E-07
	(4.42408)	(17.5858)	(0.13258)	(14.4398)	(92.0622)	(1.7E-05)
	[-1.80618]	[-1.42408]	[2.75843]	[1.11389]	[2.00983]	[-0.03678]
X13 (-1)	-0.053598	-0.248377	4.88E-06	0.662852	1.715978	3.12E-07
	(0.05471)	(0.21746)	(0.00164)	(0.17856)	(1.13840)	(2.1E-07)
	[-0.97974]	[-1.14219]	[0.00298]	[3.71231]	[1.50737]	[1.50258]
X16 (-1)	-0.010063	-0.003645	-0.000176	-0.048932	-0.430573	-1.52E-08
	(0.01207)	(0.04796)	(0.00036)	(0.03938)	(0.25107)	(4.6E-08)
	[-0.83404]	[-0.07600]	[-0.48727]	[-1.24258]	[-1.71497]	[-0.33319]
RX1 (-1)	-73,944.26	-1,065,749	-2059.836	697,560.5	-1,584,385	1.205415
	(98,805.8)	(392,755)	(2960.96)	(322,493)	(2,056,084)	(0.37446)
	[-0.74838]	[-2.71353]	[-0.69566]	[2.16303]	[-0.77058]	[3.21905]
C	24,332.33	168,739.4	406.6797	-34,794.69	-21,707.89	-0.061113
	(17,409.6)	(69,203.3)	(521.721)	(56,823.1)	(362,282)	(0.06598)
	[1.39764]	[2.43832]	[0.77950]	[-0.61233]	[-0.05992]	[-0.92624]

waterway freight index, $RX_{3(t-1)}$ is waterway freight index for $(t - 1)$ -th year, RX_4 is aviation freight index, $RX_{4(t-1)}$ is aviation freight index for $(t - 1)$ -th year.

4 Validation of Model

Through the actual data of comprehensive transportation in 2018, forecast models of freight index were verified. The actual values of rail freight index (RX_1), highway freight index (RX_2), waterway freight index (RX_3) and aviation freight index (RX_4) were shown in Table 7. The freight index of each transport mode in 2018 was predicted and compared with the actual value in 2018. The forecast results and error rate of the model were shown in Table 7.

From Table 7, the accuracy of four VAR models established in Eqs. (3) and (6) were above 90%. The accuracy of highway freight index (RX_2) and railway freight index (RX_4) was over 95%, so the models were in good agreement with the reality.

Table 7 Validation of model

Model	Actual value	Forecast value	Error rate (%)
(3)	0.078313	0.076463	-2.36231
(4)	0.769335	0.777434	1.05283
(5)	0.135833	0.126245	-7.05867
(6)	0.000144	0.000152	5.55556

5 Conclusion

The 26 indicators were selected on the basis of production, supply, benefit and macro-economy. The basic indicators were divided into 13 leading indicators and 13 synchronous indicators by cluster analysis. Factor analysis was used to analyze the importance degree of each leading indicator and the freight characteristic quantities were determined. The Granger causality test was carried out between the four freight indexes and the five characteristic quantities. The relationship model between freight index and characteristic quantities was constructed. Finally, the validity of the model was verified through comparing actual value with forecast value. The results show that the freight index is volatile and the characteristic quantity is stable. The error between forecast and actual value of freight index is less than 10%, which verifies the validity of the model. The results can provide theoretical support for forecasting the development characteristics of freight transport among various modes of transport, and provide a decision-making basis for national macro departments to make policies.

Acknowledgements This research was financially supported by Economic Operation Adjustment Bureau of National Development and Reform Commission. The authors had the gratitude to the reviewers and the doctoral students of Institute of Traffic Engineering for their valuable comments.

References

1. Kuzu S, Önder E (2014) Research into the long-run relationship between logistics development and economic growth in Turkey. *Soc Sci Electron Publ* 3(1):11–16
2. Gao Y, Zhang Y, Li H et al (2016) Study on the relationship between comprehensive transportation freight index and GDP in China. *Procedia Eng* 137:571–580
3. Guo J, Xia J (2011) Econometrical investigation on infrastructure investment and economic development in China: a case study using vector autoregression approach. *KSCE J Civ Eng* 15(3):561–567
4. Muller S, Klauenberg J, Wolfermann A et al (2015) How to translate economic activity into freight transportation. *Transp Res Procedia* 8:155–167
5. Meng L, Li L, Liu S (2015) Research on relationship between regional railway freight volume and GDP. In: *Proceedings of 4th international conference on logistics, informatics and service science, Heidelberg*, pp 345–350
6. Pandawa, Pucangan, Kartasura et al (2018) The relationship between transport, economic growth and environmental degradation for ninety countries. *J Environ Sustain* 2(1):1–64

7. Pradhan RP (2019) Investigating the causal relationship between transportation infrastructure, financial penetration and economic growth in G-20 countries. *Res Transp Econ* 78:100766
8. Lei MCQ (2014) Research prospect on the relationship between transportation and regional economy in the Hui nationality agglomeration of China. *Mod Transp* 3(4):56–61
9. Tong T, Yu TE (2018) Transportation and economic growth in China: a heterogeneous panel cointegration and causality analysis. *J Transp Geogr* 73:120–130
10. Jiang X, He X, Zhang L et al (2017) Multimodal transportation infrastructure investment and regional economic development: a structural equation modeling empirical analysis in China from 1986 to 2011. *Transp Policy* 54:43–52
11. Duru O, Yoshida S (2011) Long term freight market index and inferences. *Asian J Ship Logist* 27(3):405–421
12. Kumar A, Anbanandam R (2019) Development of social sustainability index for freight transportation system. *J Clean Prod* 210:77–92
13. Tsioumas V, Papadimitriou S, Smirlis Y et al (2017) A novel approach to forecasting the bulk freight market. *Asian J Ship Logist* 33(1):33–41
14. Duru O, Bulut E, Yoshid S (2010) Bivariate long term fuzzy time series forecasting of dry cargo freight rates. *Asian J Ship Logist* 26(2):205–223
15. Angelopoulos J, Duru O, Chlomoudis C et al (2016) Spectral dynamics of dry cargo shipping markets: theory of long waves fact or artifact? *Int J Transp Econ* 43(3):227–246
16. Khan MZ, Khan FN (2020) Estimating the demand for rail freight transport in Pakistan: a time series analysis. *J Rail Transp Plan Manag* 14:100176
17. Moscoso-López JA, Turias IJT, Come MJ et al (2016) Short-term forecasting of intermodal freight using ANNs and SVR: case of the port of Algeciras Bay. *Transp Res Procedia* 18:108–114
18. Uyar K, İlhan Ü, İlhan A (2016) Long term dry cargo freight rates forecasting by using recurrent fuzzy neural networks. *Procedia Comput Sci* 102:642–647
19. Adland R, Cariou P, Wolff FC (2017) What makes a freight market index? An empirical analysis of vessel fixtures in the offshore market. *Transp Res Part E Logist Transp Rev* 104:150–164
20. Betanzo QE, Romero JA (2010) An urban freight transport index. *Procedia Soc Behav Sci* 2(63):12–22
21. Løvold Rødseth K (2017) Productivity growth in urban freight transport: an index number approach. *Transp Policy* 56:86–95
22. Gong WW, Zhu JC, Feng SS (2016) Compilation and application of railway coal transport boom index. *Railw Freight* 34(12):19–24
23. Shi QZ, Yang XM (2000) Calculation of horizontal dynamic prosperity index of various transportation modes. *J Shanghai Jiao Tong Univ* 34(S1):5–7
24. Ning XL, Zhao YF, Chen FE (2013) Research on global shipping climate index. *J Shanghai Jiao Tong Univ* 47(03):417–422+427
25. Liu B, Jiang GY, Li H et al (2014) Scheme design and empirical research on an index system for road freight in Ningbo region. *Procedia Soc Behav Sci* 138(2):23–32

An Optimization Model of Passenger Transportation Structure Based on Land Use and Energy Consumption



Fenghe Yu, Yining Lou, Lifei Han, and Jieyu Zhu

Abstract Passenger transportation structure (PTS) features a huge effect on the proficiency of the territorial comprehensive traffic framework. Therefore, one of the important means to improve its efficiency is to rationalize and optimize the passenger transportation structure. From a static point of view, this paper puts forward user optimum thought to meet the demand of regional passenger optimal thought and constructed an optimization model of the passenger transportation structure on the basis of land and energy consumption. Through analysis of the passenger transportation systems of Harbin, Daqing, and Yichun in China, a model was designed by simulation and the proportion of the traveler transportation organization were determined. The method of fuzzy comprehensive was used to assess the traveler transportation arrangement, which is based on optimization model and real-world data. The effectiveness of the optimization model was demonstrated by analysis of the evaluation results. The displayed method and estimation results can help transportation organizers and policy-makers in deciding future traveler transportation foundation speculation needs.

Keywords Passenger transportation structure · Optimization model · Land use · Energy consumption · Regional comprehensive traffic · Fuzzy evaluation

F. Yu

Heilongjiang Chianfarm Highway and Bridge Architectural Engineering Co., Ltd, Harbin 150090, China

Y. Lou (✉)

Department of Mathematics, University College London, London WC1E 6BT, UK
e-mail: yining.lou.20@ucl.ac.uk

L. Han · J. Zhu

School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China

1 Introduction

Transportation structure refers to the mode part of entry trips (railroad, interstate, flying, etc.), which can reflect the characteristics of activity request and the useful parts of different modes of transportation. It features a coordinate effect on the arrangement of the constrained transportation assets and decides the proficiency of a regional transportation framework [1]. Passenger transportation takes up nearly 18% of the energy utilization in the US and about 63% of the transportation sector's entire energy utilization [2, 3]. Globalization, energy involvements, vehicle innovations, and progressed modeling abilities make the current period an appropriate time to return to the impact of land and energy utilization on PTS. Optimizing the arrangement of traveler transportation for sustainable improvement of regional transport has practical applications worldwide.

The impact of different variables on a PTS has been examined broadly in past findings. Population is regarded as one of the foremost critical variables affecting traveler transportation [4–6]. Various other factors have been studied and observed to impact PTS change, such as salary [7, 8], education level, and number of children [9]. In addition, fuel and trip prices moreover have been recognized as components influencing PTS [7, 10]. Land use is regarded as one of the major variables affecting PTS and other different indicators have been investigated in past study, such as density and city advancement [9, 11, 12]. The other two extra variables of road length and current capacity also affect PTS [8].

Furthermore, prior researches have explored the impact of population social and financial descriptions, salary, education, land utilization, highway capacity, and fuel costs on PTS.

The viewpoint of “transportation allocation proportion” was first proposed by the study of traffic structure in the research center of city transport in Chicago, and its purport is to attain the best combination by making use of resources according to various convenience and purposes [13]. In addition, some scholars have put forward “transfer curve method”, which has been highly applied in transportation planning practices over the last twenty years [14]. With the scholars raising the depth of research on the aggregate model, they gradually discovered its deficiency and begun to study a disaggregate model. Ben-Akiva cites a theory of economic called “Utility Theory” to build the disaggregate model which is based on Probability Theory and then he applies this model to actual [15]. A mathematical model has also been established, which decomposes variations in energy utilization to development in transportation capacity, organizational change, or modal change [16]. Traveler transportation models are coordinated inside the framework of the National Long-term Energy and Transportation Planning computer program [17]. Utilizing panel information from the 48 states amid the period of 1998–2008, a random model was built to figure the entire amount of traveler transportation based on energy utilization and emissions of greenhouse gas in the US [18].

Chinese scholars have also completed related work in traffic structure optimization. They have analyzed the rules of integrated transport development in China

and improved the foreign methods and theories [19, 20]. Others have introduced the conceptions of environmental capability, energy utilization of the methods, and transportation capability to construct the optimization model of transportation organization based on energy utilization to decide a realistic transportation organization with minimal energy utilization [21]. Some analysis of development laws and influence factors of traffic structure in China have proposed evaluation standards and confirmed goals and measures of future traffic structures [22]. Wu studied the matching of planning positioning and rail network from two aspects of the central system of the spatial axis of the metropolitan area, and found that the matching of the two is poor in the macro layout of the metropolitan area [23]. Huang uses the macro traffic software TransCAD for modeling, and analyzes the improvement effect of the traffic operation condition of Siping Road from a macro perspective [24]. Lin selected the “flow” data of railway and highway passenger transport between 20 cities in the Haixi city group and analyzed the pattern and structure of inter-city correlation network through ARCGIS and social network analysis [25]. Markov chain theory has been used to predict passenger structure evolution [26]. Finally, an improved model of the passengers’ distribution ratio in a regional transportation corridor based on multidimensional logit mode considering the passenger properties has been established [27]. Ma et al. used econometrics to explore the connection between transport and financial development for different regions. And suggestions of regional sustainable development between transport and economy structure are offered [28]. Zhang and Bai adopted best–worst methods to analyze the components hindering the vehicle fuels promotion of biodiesel and measure the sustainability of different biodiesel assessment [29].

Investigation on the advancement of traveler transportation structure based on land and energy constraints in China can back the policies improvement, which can increase the utilization of proficient transportation modes. However, it is essential to better identify the impacts of transportation, density of population, and policy factors on passenger turnover related to a transportation mode [30].

This study’s authors propose a methodology to investigate how land use and energy consumption impacts a PTS in China. Referring to prior research results, a resource consumption and optimization model was established. The first step was to estimate the quantity of PTS in various cities as a work of socioeconomics factors, fuel utilization and land use. The random parameter model for assessing whole PTS make coefficients changes available. The proposed model is verified through the case study and the fuzzy integrated assessment. The model results of PTS can provide reference of reasonable investment of energy and transportation facilities for transport policy-makers in the future.

2 Methodology

2.1 Modeling

Saving resources is one of the aims of optimizing a passenger transportation structure. Traffic infrastructure construction consumes significant resources and if facilities are not deployed properly, it may result in a huge waste of these resources [31, 32]. Equation (1) was developed to calculate the sharing rate of passenger transportation with the purpose of saving land resource and energy.

$$\min M = \sum_j (u_{Land} \rho_j + u_{energy} \gamma_j) \alpha_j V_P \tag{1}$$

$$s.t. \begin{cases} \sum_j V_j^0 \leq V_P \\ V_j^0 \leq \alpha_j V_P \leq V_j^{\max} \\ \sum_j \gamma_j \alpha_j V_P \leq E^{\max} \\ \sum_j \rho_j \alpha_j V_P \leq L^{\max} \end{cases}$$

where, M —total resources occupied by passenger transportation means (including land and energy);

u_{Land}, u_{energy} —unit price of land and energy resources;

ρ_j, γ_j —land and energy cost of unit passenger volume occupied by the passenger transportation means of j ;

α_j —share rate of the passenger transportation means j ;

V_j^0 —passenger volume of passenger transportation means j , 10,000 people per day;

V_j^{\max} —largest capacity of the passenger transportation means j , 10,000 people per day;

V_P —total passenger traffic volume in the regional, 10,000 people per day;

E^{\max} —limits of energy consumption;

L^{\max} —limit of land resources.

In order to make the two different dimension variables become unified dimensions, land and energy price were introduced. The objective function meets passenger transportation demand and, at the same time, sets consumption of land and energy resources to the minimum while other resources are bounded by their upper limits.

2.2 Instructions for Model Parameters

The approximate unit price of land and energy resources in the model, u_{Land} and u_{energy} , can be obtained from historical data. Specifically, the model can use historical

total area and total consumed energy for each means of transportation to divide the corresponding total passenger transportation volume. Then, the time series may be analyzed.

Corresponding limits of land resources can be determined by land use planning for relevant planning departments. The energy consumption limit can be estimated with the following method. The energy limit for a PTS is the variance of energies between total quantity and non-transportation-related activities. After determining the above limit and total passenger traffic volume in the region, a reasonable structure comprised of various kinds of passenger transportation means can be developed using the model.

3 Case Study

The study investigated the comfort, cost, convenience, security and time of three modes among cities of Harbin, Daqing, and Yichun in Heilongjiang Province. This paper analyzed the adjustment of the existing structure of traveler transportation for zones by utilizing the optimization model above.

3.1 Data Processing

On the basis of historical data from survey, the real passenger capacity and supply of various transportation modes among three cities was decided, as displayed in Fig. 1. H-D represents Harbin-Daqing, H-Y represents Harbin-Yichun, D-Y represents Daqing-Yichun, D-H represents Daqing-Harbin, Y-H represents Yichun-Harbin, Y-D represents Yichun-Daqing.

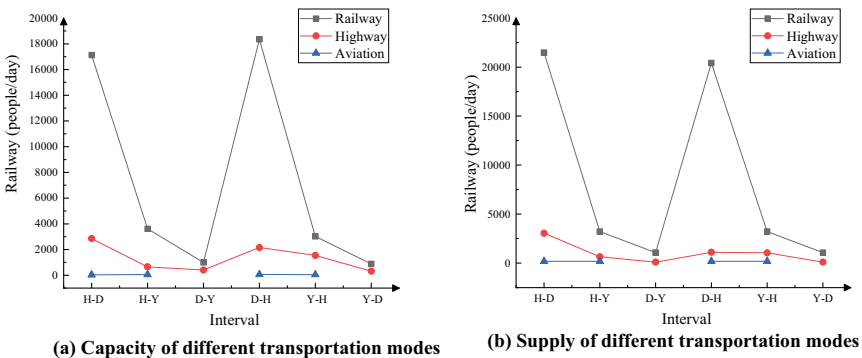


Fig. 1 The real passenger capacity and supply of various transportation modes

Table 1 Distribution rate of traveler transport

Interval	Railway	Highway	Aviation
H-D	0.361	0.613	0.026
H-Y	0.511	0.415	0.074
D-Y	0.411	0.589	–

3.2 Model Optimization

Employing the optimization toolbox in simulation and using the optimization model based on land and energy constraints, the distribution rate of traveler transport of the above cities was calculated. The results are shown in Table 1.

4 Model Evaluation

Analytic Hierarchy Process (AHP) can be applied to the assessment of PTS decisions. Fuzzy mathematics can express problems where it may be challenging for quantifiable study. Thus, combining the two methods can assess the rationality of PTS more accurately [8, 18]. Using the optimization example of PTS from Harbin to Daqing, the percentage of each transportation mode, and the relationship of pre-optimization and post-optimization are presented in Table 2; Fig. 2.

By contrasting the proportion of each transportation mode as it exists today with the optimized result, the model suggests that the proportion of railway should increase by 18.75%, highway should decrease by 9.05%, and airline usage should only slightly increase.

The fuzzy comprehensive evaluation was used according to the parameters outlined below.

- (1) Concerning the established indicator structure of traveler transport, the variable can be set as:

$$U = \{U_1, U_2, U_3, U_4\} = (\text{economy, society, ecology, technology});$$

$$U_1 = (\text{transportation cost, payback time of investing, infrastructure investing, profit margin});$$

$$U_2 = (\text{suitability with economic improvement, capacity to meet the request of transportation, comfort, resident sensitivity});$$

$$U_3 = (\text{energy consumption, occupancy rate of land resource, air pollution, noise pollution});$$

Table 2 Relationship of pre-optimization and post-optimization

H-D	Railway	Highway	Aviation
Before	0.304	0.674	0.022
After	0.361	0.613	0.026

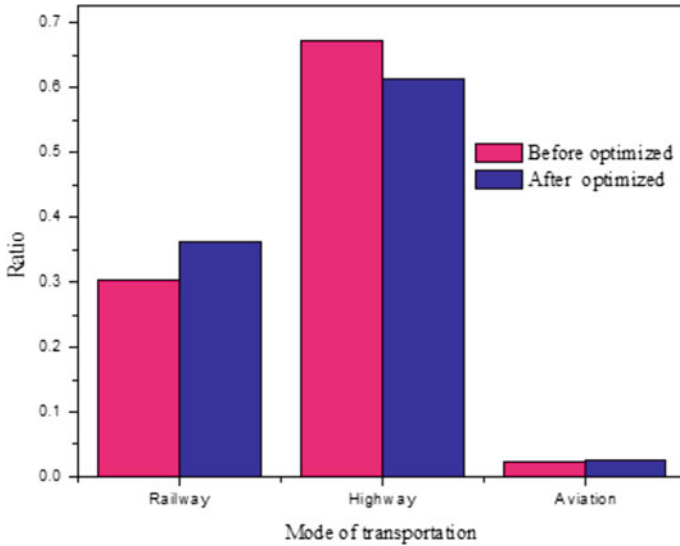


Fig. 2 Relationship of pre-optimization and post-optimization

$U_4 =$ (density of road network, speed, security, network adaptability).

- (2) Establishing an evaluation set of seven grades' = {excellent +, excellent -, good +, good -, medium, poor, very poor}.
- (3) Using AHP to get the different evaluation index weights of each layer is:
 $W = \{W_1, W_2, W_3, W_4\} = \{0.531, 0.245, 0.152, 0.072\}$.

Level analysis of highway, railway, and aviation in concrete indexes of three indicators are as shown in Table 3; Fig. 3.

According to the level value of PTS between pre-optimization and post-optimization, the weight of each indicator of pre-optimization and post-optimization can be determined as follows.

- (1) Results before optimization
 $V_1 = \{v_{11}, v_{12}, v_{13}, v_{14}\} = \{0.303, 0.125, 0.422, 0.150\}$;
 $V_2 = \{v_{21}, v_{22}, v_{23}, v_{24}\} = \{0.245, 0.315, 0.267, 0.173\}$;
 $V_3 = \{v_{31}, v_{32}, v_{33}, v_{34}\} = \{0.354, 0.221, 0.317, 0.108\}$;

Table 3 Level value of index

Means	Payback time	Capacity to meet request	Economic fitness	Comfort	Resident sensitivity	Air pollution	Speed	Safety
Highway	Long	Good	Highest	Medium	High	Serious	Medium	Low
Railway	Longest	Good	High	Good	Medium	Medium	Fast	High
Aviation	Short	Medium	Medium	Best	Low	Medium	Fastest	Highest

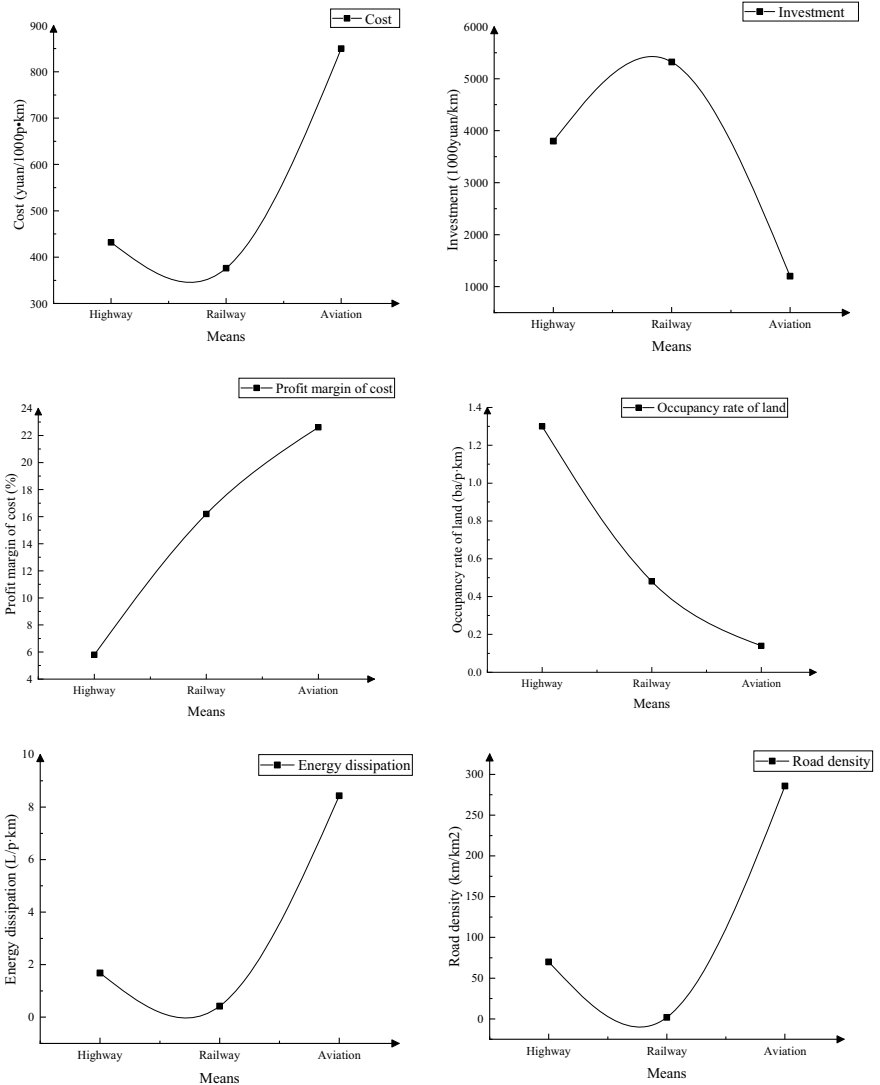


Fig. 3 Level value of index

$$V_4 = \{v_{41}, v_{42}, v_{43}, v_{44}\} = \{0.363, 0.277, 0.114, 0.246\}.$$

(2) Results after optimization

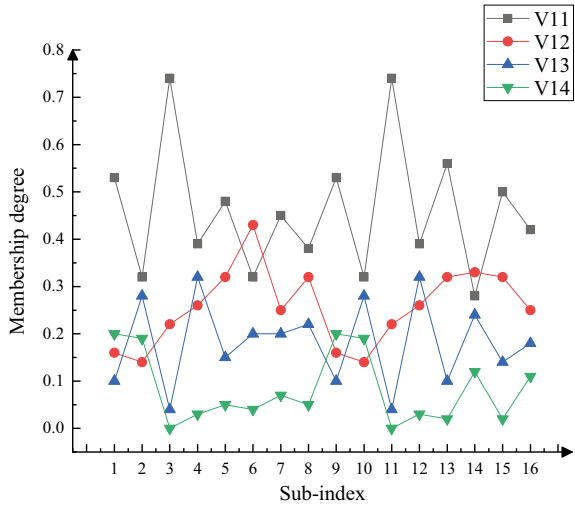
$$W_1 = \{w_{11}, w_{12}, w_{13}, w_{14}\} = \{0.324, 0.103, 0.065, 0.508\};$$

$$W_2 = \{w_{21}, w_{22}, w_{23}, w_{24}\} = \{0.335, 0.425, 0.091, 0.149\};$$

$$W_3 = \{w_{31}, w_{32}, w_{33}, w_{34}\} = \{0.466, 0.161, 0.277, 0.096\};$$

$$W_4 = \{w_{41}, w_{42}, w_{43}, w_{44}\} = \{0.314, 0.240, 0.137, 0.309\}.$$

Fig. 4 Membership degree of each factor



- (3) Using Delphi method to determine the membership degree of influence factors, as shown in Fig. 4.
- (4) Evaluation results.

According to the mean of maximum method, Harbin to Daqing passenger transport structure optimization before the “good –”, optimized for the “excellent –”, a two-level promotion.

In other words, the overall performance of the adjusted passenger structure in economy, society, ecology and technology has improved in terms of the current structure. Furthermore, the optimized structure is more responsive to the social improvement and can offer a reference for designers.

5 Conclusions and Discussions

With the goal of saving land and energy resources, this study proposed the optimization concept of least source utilization with the basis of meeting the request of zonal travelers from a static point of view. Considering factors of land and energy price, the proposed model of a PST was developed on the basis of land and energy limitations.

Based on the sample of traveler transport system in the three cities, the proposed model is computed by simulation, including data categorization, dimensionless handling and benefit function. The results show that the three cities integrated traveler transport system optimization ratio.

According to features of a PST, an indicator system for assessment of PST was established. The method of fuzzy comprehensive is used to evaluate the traveler

structure before and after optimization. Finally, the proposed model was verified by comparing the estimation results.

A regional integrated traveler transport system is a complicated system. Passenger transport structure optimization should consider many factors, including the influence of passenger travel behavior and administrative policy. Based on this study, a passenger traffic optimization model that accounts for resource constraints can provide a new perspective on the research of a regional passenger transport mode. The displayed method and assessment results can assist transportation planners and policy-makers in settling future investment of passenger transportation infrastructure.

Acknowledgements This research was financially supported by the Transportation Center funded by Ministry of Transport of the People's Republic of China (Grant number 20113188201420). The authors had the gratitude to the reviewers and the doctoral students of Institute of Traffic Engineering for their valuable comments.

References

1. Yang C, Miao ZS (2013) System dynamics model of shanghai passenger transportation structure evolution. *Procedia Soc Behav Sci* 96(6):1110–1118
2. US Department of Energy (2008) Energy, Chap 2. In: *Transportation energy data book*, 28th edn. http://cta.ornl.gov/data/edb28/Edition28_Chapter02.pdf
3. US Environmental Protection Agency (2008) US inventory of greenhouse gas emissions and sinks: 1990 to 2006. http://www.epa.gov/climatechange/emissions/downloads/08_CR.pdf
4. Fulton L, Noland B, Meszler D, Thomas J (2000) A statistical analysis of induced travel effects in the US mid-Atlantic region. *J Transp Stat* 3(1):1–14
5. National Surface Transportation Policy and Revenue Study Commission (2007) Implications of rising household income on passenger travel demand. Commission briefing paper 4A-06. Prepared by: Section 1909 Commission Staff
6. Souleyrette R, Hans Z, Garrison W, Wazny L (1995) Analysis of trends underlying urban/regional impacts of traffic growth. *J Urban Plan Dev* 158–171
7. Greene D, Chin S, Gibson R (1995) Aggregate vehicle travel forecasting model. Oak Ridge National Laboratory, Oak Ridge
8. Noland R, Cowart W (2000) Analysis of metropolitan highway capacity and the growth in vehicle miles of travel. *Transportation* 27(4):363–390
9. Brownstone D, Golob T (2009) The impact of residential density on vehicle usage and energy consumption. *J Urban Econ* 65:91–98
10. Southworth F (2001) On the potential impacts of land use change policies on automobile vehicle miles of travel. *Energy Policy* 29:1271–1283
11. Chatman D (2008) Deconstructing development density: quality, quantity and price effects on household non-work travel. *Transp Res Part A* 42:1008–1030
12. Su Q (2010) Travel demand in the US urban areas: a system dynamic panel data approach. *Transp Res Part A* 44:110–117
13. Wei JL (2005) Study of urban passenger traffic structure based on efficiency analysis. *J Traffic Manag* 5:57–60
14. Shao C (2006) *Principle of transportation planning*. China Railway Press, Beijing
15. Ben-Akiva M (1994) *Combined revealed and stated preferences data*. Springer, Netherlands
16. Tiwari P, Gulati M (2013) An analysis of trends in passenger and freight transport energy consumption in India. *Res Transp Econ* 38(1):84–90

17. Rentziou A, Gkritza K, Souleyrette RR (2012) VMT, energy consumption, and GHG emissions forecasting for passenger transportation. *Transp Res Part A Policy Pract* 46(3):487–500
18. Krishnan V, Kastrouni E, Pyrialakou VD, Gkritza K, McCalley JD (2015) An optimization model of energy and transportation systems: assessing the high-speed rail impacts in the United States. *Transp Res Part C Emerg Technol* 54:131–156
19. He DQ, Liu H, He K et al (2013) Energy use of, and CO₂ emissions from China's urban passenger transportation sector—carbon mitigation scenarios upon the transportation mode choices. *Transp Res Part A Policy Pract* 53:53–67
20. Papageorgious M (2005) An integrated control approach for traffic corridors. *Transp Res C* 3(1):19–30
21. Lu HP, Wang JW, Zhang P (2004) Optimization of urban traffic structure based on energy consumption. *J Tsinghua Univ* 44(3):383–386
22. Onden I, Guresen E (2011) Interior design of passenger coaches using fuzzy optimization. *Procedia Comput Sci* 3:491–498
23. Wu X, Wang C, Tan M (2019) Research on spatial structure optimization of Xi'an metropolitan area under the influence of rail transit. *Urban Dev Res* 26(11):21–26
24. Huang Y (2019) Traffic congestion in Siping Road and its surrounding areas can be alleviated by optimizing road network structure. *Traffic Transp* 35(06):34–36
25. Lin H (2019) Study on spatial structure of urban agglomeration in the west of China from the perspective of passenger traffic flow. In: China urban planning society, Chongqing municipal people's government. China Urban Planning Society, pp 785–794
26. Wang D, Sun NX, Li YH (2013) Using Markov process for passenger structure prediction within comprehensive transportation channel. *J Comput (Finland)* 8(4):1072–1077
27. Peng H, Han YQ et al (2010) Distribution ratio model of passenger mode in regional transportation corridor considering the passenger properties. *J Beijing Inst Technol* 19(SUPPL. 2):72–77
28. Ma Y, Zhu J, Gu G et al (2020) Freight transportation and economic growth for zones: sustainability and development strategy in China. *Sustainability* 12(24):10450
29. Zhang L, Bai W (2021) Sustainability of crop-based biodiesel for transportation in China: barrier analysis and life cycle ecological footprint calculations. *Technol Forecast Soc Change* 164:120526
30. Mishalani RG, Goel PK, Westra AM, Landgraf AJ (2014) Modeling the relationships among urban passenger travel carbon dioxide emissions, transportation demand and supply, population density, and proxy policy variables. *Transp Res Part D* 33:146–154
31. Hyman WA, Alfelor RM, Roberts PO (1997) Guidelines for multimodal corridor capacity analysis. In: Dearbon. Proceedings of the sixth TRB. AIC07.104
32. Bunker J, Ferreira L (2001) Assessing freight corridor modal performance. In: Proceedings of the third international conference on transportation and traffic studies, Guilin, pp 1–10

A Review of Motor Vehicle Driver Fatigue Research—Based on L3 Automatic Driving System



Hanying Guo and Yuhao Zhang

Abstract Fatigue driving behavior is one of the important causes of traffic accidents, and it is also an important research hotspot in the field of traffic safety. On the basis of the traditional theory of driving fatigue, this paper reviews the research status of fatigue driving behavior based on conditional autopilot system, divides driving fatigue into sleep related (SR) fatigue and task related (TR) fatigue, and describes their definitions respectively. The occurrence time and main manifestation of driving fatigue in level 3 automatic driving system are discussed. Three kinds of driving fatigue identification methods and research results are summarized, and their advantages and disadvantages and application fields are analyzed in Level 3 automatic driving system. The results show that most of the existing researches on autopilot fatigue focus on the passive fatigue caused by monotonous driving task, and few studies combined with SR fatigue; Fatigue identification methods and information source are usually unique, and the research using the fusion of multiple detection methods is not deep enough.

Keywords Fatigue driving · Automatic driving system · Fatigue classification · Recognition method

1 Introduction

In recent years, with the rapid growth of car ownership, the number of casualties in traffic accidents has increased year by year. World Health Organization statistics show that 1.2 million people die from road traffic accidents every year, and millions of people are injured in traffic accidents [1]. Among the factors causing traffic accidents, fatigue driving is one of the most dangerous driving behaviors, which is closely related to serious traffic accidents. It is one of the main causes of car rear-end accidents, which brings huge safety risks and economic burden to the society [2, 3]. Fatigue will reduce the driver's alertness, thus affecting the normal execution of

H. Guo (✉) · Y. Zhang
School of Automobile and Transportation, Xihua University, Chengdu 610039, China
e-mail: 0120070011@xhu.edu.cn

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_18

257

driving tasks, such as attention level, braking decision-making and reaction time [4–6]. Therefore, how to reduce driving fatigue has been a focus in the research field of traffic safety.

According to statistics, about 20% of traffic accidents in the world are related to fatigue driving [7]; In 2004, 90.4% of illegal drivers punished by Finland were related to fatigue [8]; A survey in Ontario, Canada, in 2008 showed that 58.6% of drivers were tired while driving, and even 14.5% of drivers reported that they dozed off while driving [9]; Of the 32,367 fatal accidents in the USA in 2012, 809 were related to fatigue driving [10].

Compared with foreign countries, China has relatively few reports on fatigue driving, but in fact, the accidental casualties caused by fatigue driving are also very serious. In 2007, there were 1768 traffic accidents related to fatigue [11]; in 2011, 1755 traffic accidents caused by fatigue driving caused 1003 deaths [12].

Because of the serious consequences caused by fatigue driving, a large number of scholars have conducted in-depth research on it. At present, research results on fatigue driving have been accumulated to a certain extent [13–19], and the research on fatigue of automatic driving is less. Some researches have explored the effect of fatigue on the driver's taking over process in the Level 3 automatic driving system. However, there is a lack of systematic review on the mechanism of driving fatigue and how to identify driving fatigue in the automatic driving system [20–24]. Considering the danger of driver fatigue, the rationality and accuracy of the method to identify and predict fatigue will greatly affect the safety and popularity of automatic driving technology. So, it is necessary to review the research results in this field and summarize the research methods of fatigue driving.

Therefore, this paper collated the relevant research results of manual driving fatigue, systematically sorted out the classification and recognition methods of driving fatigue in the existing research, discussed the advantages and disadvantages of each identification method combined with the automatic driving system, summarized the shortcomings of the existing research, and prospected the future research directions, so as to provide a basis and reference for improving the road safety of the automatic driving vehicle.

2 Classification of Driving Fatigue

For drivers, fatigue is a process of gradual accumulation, which may be caused by repeated monotonous or high-intensity driving tasks [13, 25]. May and Baldwin [26] classified driver fatigue into task-related (TR) fatigue and sleep-related (SR) fatigue in 2009. As shown in Fig. 1, this classification method is used in this paper.

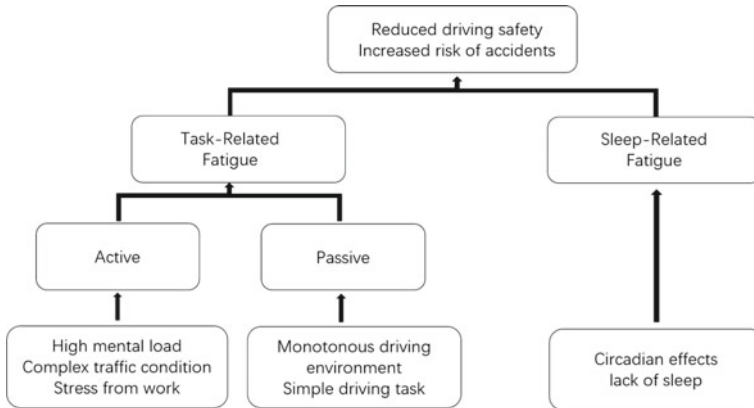


Fig. 1 Types of fatigue

2.1 Sleep-Related Fatigue

Drivers’ sleep and wakefulness follow the natural circadian rhythm of the human body. They are awake during the day and sleepy at night. When circadian rhythms alert drivers to sleep, their driving performance drops significantly. As early as 1995, Pack et al. [27] discovered that there were more traffic accidents caused by SR fatigue between 2 to 6 am and 2 pm to 4 pm than other time periods; Davenne et al. [28] investigated the driving conditions of 34 male drivers during the day and night in 2012 and found that self-reported fatigue during night driving was worse than daytime; Li et al. [10] pointed out in 2018 that young drivers have a higher frequency of fatigue driving at night than at daytime.

Lack of sleep can also affect SR fatigue. Numerous studies have shown that when sufficient sleep is not available, the driving performance of the driver will be reduced. Lemke et al. [29] evaluated the 809 professional male driver data counted by the National Sleep Foundation. The results showed that drivers with sleep disorders (8.4%) had a higher frequency of accidents due to fatigue; Kwon et al. [16] surveyed 161 professional drivers in South Korea in 2019 and found that poor sleep quality and excessive daytime sleep can cause drivers to be highly fatigued; However, the results of Jiang et al. [18] showed that the significant factor that affects driving fatigue is work pressure rather than sleep quality, they found that work pressure will force the driver to continue driving after feeling fatigue.

2.2 Task-Related Fatigue

TR fatigue is caused by high-intensity or repetitive monotonous driving tasks and driving environment. According to this definition, existing studies can be divided into two types: active fatigue and passive fatigue [30].

- (1) Active fatigue. Active fatigue is the most common TR fatigue, which is usually related to the driver's high-intensity mental load. Crizzle et al. [31] summarized 42 articles related to the health status of professional truck drivers in 2017, and found that the working pressure and sleep status of drivers are the risk factors that affect long-distance truck traffic accidents; Useche et al. [13] investigated 524 male BRT drivers in Colombia in 2017, the results show that fatigue is positively correlated with work stress, and negatively correlated with social support rate. It is concluded that high work pressure and low social support rate will lead to drivers' fatigue more easily.

In addition to the working pressure, the complex road conditions can also cause drivers' active fatigue. Liu and Wu [32] conducted a survey on 24 subjects in 2009, and the results showed that in the urban complex road environment, drivers have greater attention needs and are prone to driving fatigue; Li et al. [10] counted 38,564 car accident records from 2006 to 2011 in Guangdong Province and found that the type of road not only affected the driver's fatigue driving tendency, but also had a positive correlation with the severity of the car accident.

- (2) Passive fatigue. Passive fatigue usually occurs in highly repetitive, simple driving environments. Li et al. [33] found that driver fatigue is more likely to occur on the highway, which is a constant road, with a single environment and only a small amount of stimulation. According to the National Highway Traffic Safety Administration's statistics in 2005, 59% of fatigue behaviors occurred on multi Lane interstate highways with a speed limit of 55 miles/h, and 23% of fatigue driving behaviors occurred on 45 mile/h two lane highways. However, only 8% of the total number of fatigue driving behaviors occurred on local cities or nearby highways [34]; Farahmand and Boroujerdian [35] also reached similar conclusions after investigating 14 Iranian subjects in 2018. Their research shows that designing complex road curves in a monotonous environment can make drivers more alert.

May and Baldwin [26] found that the detection result of the driver fatigue warning system after distinguishing the cause and type of driver fatigue is more reliable in 2009; Studies have shown that [20], under the same driving time, the driver's fatigue degree under automatic driving conditions is more serious than that of manual drivers. Therefore, distinguishing the type of driver fatigue in the automatic driving system has a positive effect on improving the safety of automatic driving.

3 Driving Fatigue in L3 Automatic Driving System

With the deepening of research, automatic driving technology has been gradually put on the agenda, which has the potential to improve the safety of road traffic system [36].

Statistics show that 90% of driving accidents are caused by drivers' human errors [37], and the main purpose of the application of automatic driving technology is to improve the safety of road traffic by reducing human errors (distraction, emotional or fatigue driving, etc.) [38]. The existing research results show that with the increase of automobile automation, the traffic accidents caused by drivers' errors will be less and less [39], and the operation of automatic driving vehicles can significantly reduce traffic accidents.

According to the definition of automatic driving by American Society of Automotive Engineers, L3 automatic driving is the conditional automatic driving. The vehicle can realize automatic driving in certain traffic environments, but the driver needs to take over the vehicle under special circumstances that may occur at any time [22]. According to this definition, many scholars have studied the driving fatigue phenomenon in the L3 automatic driving system.

In 2017, German scholar Feldhütter et al. [24] found that long time conditional automatic driving will cause fatigue, and most subjects will show signs of fatigue after 15–45 min; In 2019 Vogelpohl et al. [20] found that German drivers will gradually stop paying attention to the driving environment after 15–35 min of automatic driving, and take-over time in case of emergency is longer than that of drivers who keep alert.

American scholar Saxby et al. [30] found in a survey of 168 subjects in 2013 that long-term automatic driving will reduce the driver's alertness, resulting in increased braking frequency and reduced steering response speed; German scholar Körber et al. [40] investigated the alertness of 20 participants in L3 automatic driving in 2015. The results showed that with the increase of driving time, the frequency of drivers' inattention will also increase. They believed that if the driver did not participate in driving tasks, fatigue would be induced.

Jarosch et al. [41] invited 73 participants from Garching, Germany, to conduct a 50 min conditional automatic driving experiment in 2019. The results showed that in conditional automatic driving, participants only engaged in monotonous supervision tasks would affect their vigilance and lead to delayed response to the system takeover request.

Existing studies have shown that long-term automatic driving will put drivers in a passive fatigue state that slow response and easy to make mistakes, and passive fatigue increases the incidence of traffic accidents [30]. In the L3 automatic driving system, it is the main task for drivers to keep vigilant, supervise the operation of vehicles and take over control of vehicles in emergencies. However, due to the fact that the driver does not directly participate in the driving task in the process of vehicle operation, and is in the monotonous task of supervising the vehicle, so it is easier to produce passive fatigue than manual driver.

In a word, the existing studies mainly investigate when drivers will get tired during automatic driving and what their main manifestations are. Future research should take into account the impact of the actual traffic environment. For example, in China, where mixed traffic flows dominate, the traffic density is relatively high, and the complex traffic environment forces drivers to keep vigilance at all times, and the time for drivers to concentrate their attention may be prolonged. In such a complex and changeable traffic environment for a long time, drivers need to keep a high degree of vigilance for a long time, which may lead to active fatigue. In addition, how to effectively reduce driving fatigue and ensure that drivers can take over the control of the vehicle in a timely manner is the key to achieve automatic driving safety. Prior to this, the identification method suitable for the L3 automatic driving system should be selected in the existing driving fatigue research methods.

4 Research Methods of Driving Fatigue

Since the late 1990s, driver fatigue monitoring technology has made great progress in the field of driving behavior research [42]. Various systems use different features of vehicle or human body to realize fatigue state recognition in different angles, including vehicle running characteristics, driver physiological and facial features.

The fatigue identification method based on vehicle operating characteristics aims to judge the driver's fatigue status by monitoring the vehicle operating data, such as the steering wheel correction frequency and lane departure [43]. This kind of measurement of driving performance can get the feedback that whether the driver is paying attention while driving, but it cannot directly judge whether the driver is in a state of fatigue. Therefore, many scholars have proposed methods to detect driver's physiological signals and facial features to judge fatigue status, such as electroencephalogram (EEG), electrocardiogram (ECG) and facial movements [15, 30, 44].

As shown in Table 1, this paper reviews the existing research methods of driving fatigue from the aspects of vehicle running characteristics, driver physiological characteristics and driver facial features.

4.1 *Based on Vehicle Operating Characteristics*

The driver's fatigue state can be directly reflected in the characteristics of vehicle operation. The existing researches mainly focus on the changes of lateral motion amplitude caused by driving fatigue, the control state of steering wheel and car following behavior.

Sayed et al. [45] found in 2001 that driving fatigue will lead to a greater range of vehicle lateral motion, and the driver fatigue state can be judged by the change of vehicle running trajectory.

Table 1 Research methods of driving fatigue

Detection method		Literature number	Years
Based on vehicle operating characteristics	Lateral trajectory	[45]	2001
		[46]	2009
		[47]	2017
		[48]	2017
	Longitudinal trajectory	[2]	2016
		[49]	2019
Based on physiological characteristics	EEG	[50]	2011
		[51]	2017
		[44]	2018
	ECG	[52]	2018
		[21]	2020
		[53]	2013
Based on facial features		[54]	2014
		[18]	2016
		[55]	2016
		[56]	2020
		[57]	2020

Feng et al. [46] installed an angle sensor on the steering knuckle of the car in 2009 to obtain the control situation of 9 subjects under fatigue, and found that the frequency of steering wheel correction for fatigued drivers will be reduced; Zhenhai et al. [47] detected the driver fatigue state based on the steering wheel angular velocity time series in 2017; Li et al. [48] developed a fatigue recognition system based on the steering wheel angle data and the vehicle’s lateral position and yaw angle in 2017, with a recognition accuracy of 88%.

Hui et al. [2] studied the effect of fatigue on car following behavior in 2016, and the results showed that higher fatigue level would lead to the reduction of the headway between the driver and the vehicle in front; Mollicone et al. [49] obtained the driver’s emergency braking frequency based on vehicle speed data in 2019 and found that as the degree of fatigue increases, the driver’s hard braking frequency increases significantly.

4.2 Based on Physiological Characteristics

There are many physiological characteristics of the human body that are highly related to fatigue, such as EEG and ECG. Through the real-time monitoring of these physiological characteristics data, the driver’s fatigue status can be timely feedback,

especially EEG. Research shows that a series of physical and mental activities related to driving will be reflected in EEG.

Simon et al. [50] found in 2011 that compared with EEG alpha band power, the alpha spindle in EEG signal improves the sensitivity of fatigue detection and can better evaluate driver fatigue; Kong et al. [51] analyzed the fatigue status of drivers by collecting the instantaneous phase synchronization characteristics of EEG signals in 2017. The results show that most neurons in the cerebral cortex will discharge synchronously when fatigue occurs; Guo et al. [44] collected EEG signals from the posterior brain region of the driver when operating the driving simulator in 2018, and judged whether the driver was distracted by the driving environment.

The ECG can also reflect the driver's fatigue state. Huang et al. [52] pointed out in 2018 that wearable ECG recorders can only accurately predict fatigue status by using only 2–3 heart rate indicators; Jiang et al. real-time monitoring the driver's heart rate fluctuations in 2020 and found that the increase in driver fatigue in low-pressure and hypoxic plateau environments is accompanied by a decrease in heart rate and heart rate fluctuation amplitude.

4.3 Based on Facial Features

The visual features of the driver's facial organs can intuitively reflect the fatigue state. This method first locates the face through the geometric features of the face, and then extracts the visual features of the face to judge the fatigue state. Therefore, locating faces quickly and accurately is the basic task of detecting driver fatigue. A large number of studies have discussed methods to optimize face recognition.

Chang et al. [54] extracted the driver's facial fatigue features by fusing the global independent discrete cosine transform features and local dynamic Gabor features in 2013; Lei et al. [55] proposed a side face contour extraction algorithm in 2016 to extract the contour lines of the driver's nose and chin to detect the nodding behavior caused by fatigue; Li et al. [56] adopted improved YOLOv3-tiny convolutional neural network to capture the face area under complex driving conditions in 2020, and proposed a face feature vector to detect the fatigue state by constructing a face feature triangle, with recognition accuracy up to 94%.

Chang and Chen [54] used infrared cameras to extract the driver's face and eye positions in 2014, and then used the average eye closure speed (AECS) to identify driving fatigue levels. Under normal circumstances, the recognition success rate can reach 97.8%, but when the driver wears glasses, the success rate will be reduced to 84.8%.

Jiang et al. [18] proposed an algorithm based on image topology analysis technology, Haar like features and extreme learning machine in 2016, which can judge the fatigue state by grasping the eye movement track of the driver; Du et al. [57] integrated the three characteristics of heart rate, degree of eye opening and degree of mouth opening in 2020, and proposed a multimodal fusion recurrent neural network model, which can accurately identify fatigue state in complex driving environment.

In summary, the identification method based on vehicle operating characteristics determines driver fatigue by collecting vehicle operating data. It is non-invasive and does not cause too much interference to the driver, but driving fatigue is not the only factor that makes drivers change the running characteristics of vehicles, and it may also be emotional driving. Therefore, such methods are mostly used to analyze the causes of traffic accidents.

The identification method based on physiological characteristics is currently the most reliable and accurate fatigue index. They directly reflect the driver's physical activity, but this method requires the driver to wear signal monitoring equipment (EEG and ECG monitoring require sensors to be installed on the driver's head and chest respectively), which will cause serious interference to the driver, and is not suitable for using in real driving environment. Although many studies have optimized this kind of fatigue monitoring equipment from portability or low interference [14, 58, 59], compared with other methods, there are still many deficiencies in reducing the interference to drivers.

The identification method based on facial features is the least intrusive, but the required recognition algorithm is complicated, and the recognition accuracy is greatly affected by the driving environment and the driver's facial decorations.

It can be seen that in most current studies, only a single information source is used to detect fatigue, there are many limitations mentioned above, and there are obvious shortcomings in reliability. Therefore, in future research, the limitation of a single information source should be broken through, and more identification methods should be integrated for fatigue detection. For example, when the driver wears glasses, only using facial feature recognition will lead to a reduction in recognition accuracy, but a recognition method that combines facial features and physiological features can determine whether the driver is fatigued from different angles and improve the anti-interference of the recognition system.

4.4 Discussion

Studies have shown that the application of driver fatigue detection technology can effectively avoid traffic accidents caused by fatigue driving. The successful application of this technology depends on the cause and type of fatigue [26].

In the discussion of the types of driving fatigue, existing studies dividing driving fatigue into SR fatigue and TR fatigue, among which TR fatigue can be divided into active fatigue and passive fatigue. This paper does not explore the influence of driver's age on driving fatigue. Although many studies have shown that age factors do affect driver's fatigue state [17, 18, 60], some studies show that the influence is too complex [10], so this paper will not discuss it.

It has been previously discussed that the fatigue behavior in the L3 automatic driving system is almost all passive fatigue. This fatigue will cause the driver to have a delayed reaction when he needs to take over the control of the vehicle, which can

easily lead to accident. However, most of the current studies require subjects to have a good sleep the day before, which ignores the effects of SR fatigue.

In view of this, the driver fatigue detected in future research should be further divided into SR fatigue and TR fatigue, and corresponding countermeasures should be put forward.

For example, when the driver SR fatigue is detected, the vehicle stops running and indicates that the driver needs to rest; When detecting TR passive fatigue, it is appropriate to play a voice reminder about the surrounding driving environment to the driver, or propose a non-driving-related task (NDRT) to make the driver properly distracted to deal with other tasks [23, 24, 41], which can effectively alleviate their fatigue and improve their alertness during highly automated driving period.

At present, there are three main research methods of driving fatigue. The identification method based on vehicle operation characteristics is the least invasive method. However, for L3 automatic driving system, the vehicle runs in a specific environment. Therefore, the fatigue condition of drivers in such driving systems is hardly reflected the operation characteristics of vehicles. Although the fatigue state of driver can be manifested in characteristics such as takeover time and braking frequency, the fatigue monitoring system cannot warn the driver based on these characteristics because the vehicle has approached danger after detecting an extended takeover time or an increase in braking frequency.

The other two types of methods also have a limited role in highly automated driving (Level 3 and above). They prefer to monitor the driver's fatigue status in real time, but after the monitoring system detected fatigue driving phenomena and gave instructions, the driver has often entered fatigue threshold [42]. This means that when the automated driving system requires a driver to take over control, the driver is likely to be unable to make timely feedback due to fatigue.

Therefore, in this case, what we need is a monitoring system that can predict the fatigue state of the driver. Once the driver reaches a certain fatigue state, it will give instructions to alert the driver so as to provide enough time for the driver to take over the control of the vehicle when the automatic driving system reaches its limitation.

5 Summary and Outlook

At present, driving fatigue is the main cause of traffic accidents, and it is also a major problem that hinders the popularization of autonomous driving technology. Therefore, it is necessary to study the types of fatigue that occur under different driving conditions, to explore how different types of fatigue affect the driver's information processing and decision-making process, and find out methods and means that can effectively identify driving fatigue.

This article summarizes 59 studies on driving fatigue behavior and draws the following conclusions:

- (1) Driving fatigue can be divided into SR fatigue and TR fatigue, and TR fatigue is divided into active fatigue and passive fatigue. SR fatigue is caused by the driver's circadian rhythm phenomenon and lack of sleep, and high-intensity and repetitive monotonous driving tasks and driving environment lead to active fatigue and passive fatigue, respectively.
- (2) In the L3 automatic driving system, the main task of the driver is to supervise the operation of vehicles and to take over control of vehicles in emergency. This monotonous driving environment can easily lead to passive fatigue, cause a delay in driving response and extend the time to take over the vehicle. Although existing studies have clearly defined that a monotonous driving environment will lead to passive fatigue, it has not fully considered the impact of the complex traffic environment. In the future, it is necessary to consider the impact of current actual traffic environment in China on the driver's fatigue state, and further clarify the types of driver's fatigue in the actual traffic environment.
- (3) Currently, there are three main methods of fatigue identification: the first is the vehicle operating characteristics, although it does not affect the normal driving of the driver, they cannot reflect the fatigue state of the driver in the L3 automatic driving system; the second is the physiological characteristics of the driver, it is independent of the driving environment, accurately and directly reflect the real-time fatigue state of the driver, and can predict the occurrence of fatigue to a certain extent. However, because of its strong invasiveness, it will seriously affect the driver's driving experience, so it is mostly used in theoretical research; The last is the facial features, the identification method is greatly affected by driving environment, and the recognition algorithm is complex, but it has the advantages of non-invasive and high identification accuracy, which makes this method become a hot spot in current research. The future research should avoid using a single information source to monitor fatigue, and combine multiple fatigue identification methods to perform fatigue detection and monitoring from different perspectives, so as to avoid the shortcomings of single identification method to the greatest extent and improve the accuracy of fatigue identification.

Acknowledgements Supported by: The National Natural Science Foundation of China (No. 61803314)

Youth Scholars Foundation of Xihua University (No. DC1900007166)

References

1. World Health Organization (2015) Global status report on road safety. World Health Organization, Department of Violence and Injury Prevention and Disability, Geneva
2. Hui Z, Wu C, Yan X, Qiu TZ (2016) The effect of fatigue driving on car following behavior. *Transport Res Part F Traff Psychol Behav* **43**:80–89

3. Kocakülah G, Algül G, Köysal O (2019) constructing multi-scale entropy based on the empirical mode fatigue decomposition (EMD) and Its application in recognizing driving. *J Mol Liquids* 112182
4. Fountas G, Pantangi SS, Hulme KF, Anastasopoulos PC (2019) The effects of driver fatigue, gender, and distracted driving on perceived and observed aggressive driving behavior: a correlated grouped random parameters bivariate probit approach. *Anal Methods Accident Res* 22:100091
5. Chen J, Wang H, Wang Q, Hua C (2019) Exploring the fatigue affecting electroencephalography based functional brain networks during real driving in Young males. *Neuropsychologia* 129:200–211
6. Anderson C, Horne JA (2013) Driving drowsy also worsens driver distraction. *Sleep Med* 14(5):466–468
7. Dong N et al (2019) A WPCA-based method for detecting fatigue driving from EEG-based internet of vehicles system. *IEEE Access* 7:124702–124711
8. Radun I, Radun JE (2009) Convicted of fatigued driving: who, why and how? *Accid Anal Prev* 41(4):869–875
9. Vanlaar W, Simpson H, Mayhew D, Robertson R (2008) Fatigued and drowsy driving: a survey of attitudes, opinions and behaviors. *J Safety Res* 39(3):303–309
10. Li Y, Yamamoto T, Zhang G (2018) The effect of fatigue driving on injury severity considering the endogeneity. *J Safety Res* 64:11–19
11. Wang X, Xu C (2016) Driver drowsiness detection based on non-intrusive metrics considering individual specifics. *Accident Anal Prevent* 95:350–357
12. Traffic Management Bureau, Ministry of Public Security, PRC (2011) China Road Traffic Accidents Annual Statistical Report
13. Useche SA, Ortiz VG, Cendales BE (2017) Stress-related psychosocial factors at work, fatigue, and risky driving behavior in bus rapid transport (BRT) drivers. *Accid Anal Prev* 104(April):106–114
14. Zhao Y et al (2018) A portable high-density absolute-measure nirs imager for detecting prefrontal lobe activity under fatigue driving. *Microelectron Reliab* 82:197–203
15. Zheng C, Xiaojuan B, Yu W (2016) Fatigue driving detection based on Haar feature and extreme learning machine. *J China Univ Posts Telecommun* 23(4):91–100
16. Kwon S, Kim H, Kim GS, Cho E (2019) Fatigue and poor sleep are associated with driving risk among korean occupational drivers. *J Transp Health* 14:100572
17. Meng F et al (2019) Temporal patterns of driving fatigue and driving performance among male taxi drivers in Hong Kong: a driving simulator approach. *Accident Anal Prevent* 125:7–13
18. Jiang K et al (2017) Why do drivers continue driving while fatigued? An application of the theory of planned behaviour. *Transp Res Part A Policy Pract* 98:141–149
19. Jing D, Zhang S, Guo Z (2020) Fatigue driving detection method for low-voltage and Hypoxia Plateau Area: a physiological characteristic analysis approach. *Int J Transp Sci Technol* 9(2):148–158
20. Vogelpohl T, Kühn M, Hummel T, Vollrath M (2019) Asleep at the automated wheel—sleepiness and fatigue during highly automated driving. *Accid Anal Prev* 126(March):70–84
21. Feldhutter A, Ruhl A, Feierle A, Bengler K (2019) The effect of fatigue on take-over performance in urgent situations in conditionally automated driving. In: 2019 IEEE intelligent transportation systems conference, ITSC 2019, vol 15392, pp 1889–1894
22. McKerral A, Boyce N, Pammer K (2019) Supervising the self-driving car: situation awareness and fatigue during automated driving. In: *Adjunct Proceedings—11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019*, pp 315–20
23. Jarosch O, Kuhnt M, Paradies S, Bengler K (2017) It's out of our hands now! effects of non-driving related tasks during highly automated driving on drivers' fatigue. In: *Driving Assessment Conference*, pp 319–25
24. Feldhütter A, Hecht T, Kalb L, Bengler K (2019) Effect of prolonged periods of conditionally automated driving on the development of fatigue: with and without non-driving-related activities. *Cogn Technol Work* 21(1):33–40

25. Jamson H, Merat N, Carsten O, Lai F (2011) Fully-automated driving: the road to future vehicles. In: *Driving Assessment Conference*, pp 2–9
26. May JF, Baldwin CL (2009) Driver fatigue: the importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transp Res F Traffic Psychol Behav* 12(3):218–224
27. Pack AI et al (1995) Characteristics of crashes attributed to the driver having fallen asleep. *Accid Anal Prev* 27(6):769–775
28. Davenne D et al (2012) Reliability of simulator driving tool for evaluation of sleepiness, fatigue and driving performance. *Accid Anal Prev* 45:677–682
29. Lemke MK et al (2017) Work and Sleep among Transport Operators: Disparities and Implications for Safety. *J Transp Health* 7(August):298–309
30. Saxby DJ et al (2013) Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *J Exp Psychol Appl* 19(4):287–300
31. Crizzle AM et al (2017) Health and wellness of long-haul truck and bus drivers: a systematic literature review and directions for future research. *J Transp Health* 7:90–109
32. Liu YC, Wu TJ (2009) Fatigued driver's driving behavior and cognitive task performance: effects of road environments and road environment changes. *Saf Sci* 47(8):1083–1089
33. Li D, Liu Q, Yuan W, Liu H (2010) Relationship between fatigue driving and traffic accident. *Transp Eng* 02:104–109
34. Zhang G, Yau KKW, Zhang X, Li Y (2016) Traffic accidents involving fatigue driving and their extent of casualties. *Accid Anal Prev* 87:34–42
35. Farahmand B, Boroujerdian AM (2018) Effect of road geometry on driver fatigue in monotonous environments: a simulator study. *Transport Res F: Traffic Psychol Behav* 58:640–651
36. Kyriakidis M, Happee R, De Winter JCF (2015) Public opinion on automated driving: results of an international questionnaire among 5000 respondents. *Transport Res F: Traffic Psychol Behav* 32:127–140
37. Boggs AM, Wali B, Khattak AJ (2020) Exploratory analysis of automated vehicle crashes in California: a text analytics and hierarchical Bayesian heterogeneity-based approach. *Accident Anal Prevent* 135:105354
38. Brummelen V, Jessica MO, Gruyer D, Najjaran H (2018) Autonomous vehicle perception: the technology of today and tomorrow. *Transp Res Part C Emerg Technol* 89:384–406
39. Fagnant DJ, Kockelman K (2015) Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transp Res Part A Policy Pract* 77:167–181
40. Körber M, Cingel A, Zimmermann M, Bengler K (2015) Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manuf* 3(Ahfe): 2403–2409.
41. Jarosch O, Bellem H, Bengler K (2019) Effects of task-induced fatigue in prolonged conditional automated driving. *Hum Factors* 61(7):1186–1199
42. Zhou F et al (2020) Driver fatigue transition prediction in highly automated driving using physiological features. *Exp Syst Appl* 147:113204
43. Koesdwiady A, Soua R, Karray F, Kamel MS (2017) Recent trends in driver safety monitoring systems: state of the art and challenges. *IEEE Trans Veh Technol* 66(6):4550–4563
44. Guo Z et al (2018) Detection of driver vigilance level using eeg signals and driving contexts. *IEEE Trans Reliab* 67(1):370–380
45. Sayed R, Eskandarian A (2001) Unobtrusive drowsiness detection by neural network learning of driver steering. *Proc Instit Mech Eng Part D J Automob Eng* 215(9):969–975
46. Feng R, Zhang G, Cheng B (2009) An On-board system for detecting driver drowsiness based on multi-sensor data fusion using Dempster-shafer theory. In: *Proceedings of the 2009 IEEE International Conference on Networking, Sensing and Control, ICNSC 2009*, pp 897–902
47. Zhenhai G et al (2017) Driver drowsiness detection based on time series analysis of steering wheel angular velocity. In: *Proceedings—9th international conference on measuring technology and mechatronics automation, ICMTMA 2017*, pp 99–101
48. Li Z, Chen L, Peng J, Wu Y (2017) Automatic detection of driver fatigue using driving operation information for transportation safety. *Sensors (Switzerland)* 17(6)

49. Mollicone D et al (2019) Predicting performance and safety based on driver fatigue. *Accid Anal Prevent* 126(February):142–145
50. Simon M et al (2011) Eeg alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clin Neurophysiol* 122(6):1168–1178
51. Kong W et al (2017) Assessment of driving fatigue based on intra/inter-region phase synchronization. *Neurocomputing* 219:474–482
52. Huang S, Li J, Zhang P, Zhang W (2018) Detection of mental fatigue state with wearable ECG devices. *Int J Med Informatics* 119(August):39–46
53. Sun YF, Lu B, Wang LC (2013) Dynamic facial fatigue recognition based on independent features fusion. *J Beijing Univ Technol* 39(1):63–69
54. Chang TH, Chen YR (2014) Driver fatigue surveillance via eye detection. In: 2014 17th IEEE international conference on intelligent transportation systems, ITSC 2014, pp 366–371
55. Lei J et al (2017) Side face contour extraction algorithm for driving fatigue detection. In: 2016 international conference on electrical engineering and automation, pp 1–5
56. Li K, Gong Y, Ren Z (2020) A fatigue driving detection algorithm based on facial multi-feature fusion. *IEEE Access* 8:101244–101259
57. Du G et al (2020) Vision-based fatigue driving recognition method integrating heart rate and facial features. *IEEE Trans Intell Transp Syst* 1–12
58. Lin JS, Jiang ZY (2017) An EEG-based BCI system to facial action recognition. *Wireless Pers Commun* 94(3):1579–1593
59. de Oliveira SMS et al (2018) Electroencephalographic changes using virtual reality program: technical note. *Neurol Res* 40(3):160–65
60. Obst P, Armstrong K, Smith S, Banks T (2011) Age and gender comparisons of driving while sleepy: behaviours and risk perceptions. *Transport Res F Traffic Psychol Behav* 14(6):539–542

Analysis of Compliance Between Cape Town Agreement and China's Fishing Vessel Standards



Majing Lan, Chaoyu Ruan, and Muqiu Zhou

Abstract Based on the comparative analysis between Cape Town Agreement and Chinese fishing vessel standards, this paper finds out the existing gap and puts forward suggestions on the formulation and revision of relevant standards and the standardization of fishing vessels.

Keywords Cape Town Agreement · Fishing vessels · Standards

1 Introduction

The International Convention on the safety of fishing vessels is the only Convention on the safety of fishing vessels formulated by the International Maritime Organization (IMO) so far, which puts forward higher requirements on the structure and equipment of fishing vessels [1]. China is the largest fishing vessel country in the world. As of 2015, China had 1,042,500 fishing vessels with a total tonnage of 10,878,600 tons. However, for a long time, China's fishing vessels have been showing the characteristics of low technical level, high ship age and great hidden dangers, so there is still a certain gap to meet the provisions of the Convention [2].

On November 23, 2019, 48 countries signed a public statement expressing their determination to ratify the 2012 Cape Town Agreement on fishing vessel safety. The Torremolinos declaration aims to ensure that the 2012 Cape Town Agreement enters into force before the tenth anniversary of its adoption (October 11, 2022). The countries that signed the declaration are: Argentina, Bangladesh, Belgium, Belize, Central African Republic, Chile, China, Congo, Cook Islands, Costa Rica, Croatia, Democratic Republic of the Congo, Denmark, Ecuador, Fiji, Finland, France, Gabon, Germany, Ghana, Guinea, Guinea, Guinea, Iceland, Indonesia and Ireland Kiribati,

M. Lan (✉) · C. Ruan · M. Zhou
China Waterborne Transport Research Institute, Beijing 100088, China
e-mail: lanmajing@wti.ac.cn

C. Ruan
e-mail: ruanchaoyu@wti.ac.cn

Lebanon, Liberia, Marshall Islands, Mozambique, Namibia, Holland, New Zealand, Nicaragua, Nigeria, Norway, Panama, Papua New Guinea, Peru, Papua, Sao Tome and Principe, Sierra Leone, South Africa, Spain, Togo, Uganda, the United Kingdom and Vanuatu.

Once the Convention comes into force, it will be a great challenge for China's fisheries. In the process of continuous innovation and development of fishing vessels in China, the standardization work has played a positive role in promoting. However, with the continuous improvement of the technical level of fishing vessels in China, the fishing vessel standards formulated and revised in the past have been aging with the passage of time and lost their use value. Therefore, it is necessary to compare the requirements of international conventions with the existing technical standards of fishing vessels in China, and combine with the actual situation of China to transform the relevant technical requirements of international conventions into the standards of fishing vessels in China.

2 Main Contents

The Cape Town Agreement puts forward the basic technical requirements for ensuring the safety of fishing boats, including Chapter I General Provisions, Chapter II structure, watertight integrity and equipment, Chapter III stability and related airworthiness, Chapter IV mechanical and electrical equipment and regularly unattended machinery space, Chapter V fire prevention, fire detection, fire fighting and fighting, Chapter VI crew protection, Chapter VII life-saving equipment and devices, Chapter VIII emergency Procedures, muster and drill; Chapter IX radio communication; Chapter X onboard navigation equipment and devices. The rules are a series of international standards aimed at regulating the design and construction standards of fishing vessels, ensuring the safety of fishing vessels and crew members, and striving to make all countries with fishing vessels comply with them.

3 Analysis of the Compliance Between the Convention and the Standards of Chinese Fishing Vessels

There is a big gap between the overall safety technology of fishing vessels in China and the requirements of the International Convention on the safety of fishing vessels. This paper analyzes the compliance between the current fishing vessel standards and the International Convention on the safety of fishing vessels, finds out the gap between the standards and the conventions, and puts forward the following suggestions for the formulation and revision of the relevant standards for fishing vessels in China according to the actual situation.

1. According to the Convention, the scuttle with a depth of more than 300 mm shall be provided with a crossbar or other suitable protective devices with a spacing of no more than 230 mm. When the outlet valve size in the current standard is 610 mm × 460 mm, the spacing between bars is 250 mm. It is suggested to modify the spacing of outlet gate bars in the standard.
2. Chapter 7 of the Convention stipulates that every ship with a ship length of 45 m or more should be equipped with an approved rope throwing equipment which meets the requirements of the Convention, and has requirements on the performance and installation of the rope throwing equipment. The current standard lacks the number of ropes and the waterproof requirements for the container placed by the rope thrower. It is suggested that the standard be revised.
3. The Convention has corresponding technical requirements for self-supporting semi enclosed lifeboats, rigid liferafts and rescue boats for fishing vessels. It is suggested that the standards of “technical requirements for self-supporting lifeboats for fishing vessels”, “technical requirements for rigid liferafts for fishing vessels” and “technical conditions for rescue boats” should be quoted and formulated respectively.
4. The Convention stipulates the corresponding performance standards for radio equipment on fishing vessels, and quotes 16 resolutions adopted by IMO Congress and Maritime Safety Committee. It is suggested that it should be transformed into the corresponding “fishing vessel radio equipment performance” standard.
5. Chapter 8 of the Convention includes emergency procedures, musters and drills, which specifies in detail the general emergency procedures for fishing vessels, abandonment training and participants, frequency and content of drills, performance requirements of emergency alarm system and contents of contingency deployment table. At present, there is no corresponding content in fishing vessel and ship industry standards and fishing vessel technical regulations. It is suggested to formulate the corresponding standards of “fishing vessel emergency procedure, assembly and exercise”.

4 Suggestions on Fishing Vessel Standardization

Based on the analysis of the contents of the Convention and the standards of fishing vessels in China, the following suggestions are put forward for the standardization of fishing vessels in China.

(1) Improve fishing vessel standard system

The International Convention on the safety of fishing vessels has detailed provisions on fire fighting and life-saving, crew protection, emergency procedures, assembly and exercise, radio communication and navigation equipment. However, the main contents of the current fishing vessel standard system are focused on fishing vessel hull, engine, electrical, repair and FRP fishing

vessels. It is suggested that the standard system be supplemented and improved, Increase the safety of fishing vessels.

- (2) Follow up international conventions and standards, establish and improve relevant laws and regulations

The construction of China's existing fishery laws and regulations and management system is still lagging behind the requirements of accelerating the construction of modern fisheries. Although China's existing fishing vessel safety management laws are complete, the update speed is far behind the new requirements of relevant international conventions on fishing vessel safety. Therefore, the relevant requirements of the Convention protocol should be closely followed and studied, In the process of law enforcement, it is in a favorable position to update laws and regulations.

References

1. Zhang R, Han J, Ma Y (2012) Revision progress of international fishing vessel safety convention and China's response. *World Shipping* 35(01):31–33 + 38
2. Jie Y, Yuqing R, Yongkai W, Peng S (2016) Research on the strategy of China's ocean going fishing vessels in response to the Cape Town agreement. *China Fish. Econ.* 34(04):4–9

Analysis of Non-motor Vehicle Traffic Flows and Traffic Conflicts on Isolated Bicycle Lane



Lin Wang, Jin Ran, Zhonghua Wei, Shi Qiu, and Shaofan Wang

Abstract In recent years, non-motorized vehicle (NMV) groups for the purpose of avoiding traffic jams and facilitating travel have developed rapidly in urban traffic. They share limited right-of-way on urban roads, which leads to more conflicts between each other and increasing accident risks. In this research, the traffic flow characteristics of three types of NMV, namely bicycles (B), electric bicycles (EB), and electric tricycles (ET), on three independent bicycle lanes in Beijing were extracted by two-dimensional trajectory tracking method. The relationship between different traffic conflicts and traffic density and anti-riding rate is studied. Results indicate two remarkable findings: (1) densities of B, EB, and ET, which are mostly influenced by speed, are correlated with different types of conflicts; (2) the EB conflict shows a strong linear correlation with EB density. The linear prediction model of opposing conflict is established using reverse riding ratios of B, EB, and ET. This research recommends that reverse riding of NMV should be eliminated to reduce the bicycle lane conflict. And unidirectional non-motorized roads with physical isolation are proposed for construction in urban areas with large non-motorized traffic flow. Moreover, the travel speed of NMV on bicycle lane should be further regulated to provide safe and friendly riding environment for riders.

Keywords Conflict · Non-motorized vehicle · Bicycle lane · Traffic density · Reverse riding ratio

L. Wang (✉) · J. Ran
Shandong Transportation Research Institute, Jinan 250100, China
e-mail: wangl@sdjtky.cn

Z. Wei · S. Qiu · S. Wang
College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China

1 Introduction

In recent years, with the rise and development of sharing bicycles, the application of electric bicycles in commute and food delivery, and electric tricycles for last mile freight and package delivery, the number of Non-motorized vehicles (NMV) has been increasing. NMV partly avoid emerging traffic problems caused by the urban motorization and have received extensive attentions. The trend of bicycle reinvigorating in China leads to the expanding of NMV traffic flow. Especially the floating sharing bicycle originates and rapidly develops in China since 2015. Over 30 bicycle sharing programs launched in China by the end of 2018 [1]. Big data from the Internet shows that the number of sharing bicycle users in China exceeded 270 million in 2019, and the market gradually stabilized. This phenomenon has greatly stimulated the switch of private cars to bicycles for short-distance trips. Electric bicycle is the acronym of electric-motor-powered bicycle, which has been developed for a variety of purposes over a decade ago [2]. Electric bicycle attracts much attention with the advantages in efficient, fast, and flexible. Electric bicycles used for food delivery shuttle in Chinese urban bicycle lanes with the delivery boxes installed behind the bicycle saddles. And the rising of electric bicycle increases the accident risk [3, 4]. China is the largest international marketplace for electric bikes [5], and the electric bicycle always is divided into Pedelec and E-Bike [6]. Local deliveries have been rapidly developing due to the development of e-commerce and intra-urban express delivery. Three-wheel NMV is universally used for local delivery in China [7], especially the electric tricycle. Electric tricycle is widely applied by local contribution in central city with its transport performance, flexibility, which perfectly adapts to direct-to-door deliveries.

To sum up, this study will focus on security issues of typical three types NMV, namely bicycle (B), electric bicycle (EB), and electric tricycle (ET). The technical features of these three types NMV are normalized by different standards or specifications in China, as shown in Table 1.

- **Bicycle (B)**

In China, the maximum length and width of the bicycle are 190 cm and 60 cm, respectively. Urban Road Traffic Planning and Design Code (GB 502220-95) specifies that bicycle speed is 11–15 km/h [8].




- **Electric bicycle (EB)**

The width of electric bicycle is 60–70 cm, and the width is 60–70 cm. Electric Bicycle General Technical Conditions (GB17761-1999) [9] regulates that electric bicycle speed should not be greater than 20 km/h and should travel on bicycle lane.

- **Electric tricycle (ET)**

Express Special Electric Tricycle Technical Requirements, the national mandatory standards promulgated and implemented by State Post Bureau of China on April 8, 2016, has pointed out that the electric tricycle dimensions: length of 300 cm, width of 100 cm, height of 140 cm, and load mass less than 200 kg [10]. Its

Table 1 Three types of NMV

Type	Size	Speed limit	Example
B	Length: <190 cm Width: <60 cm Height: 90–110 cm	<15 km/h	Ordinary bicycle 
EB	Length: 180–200 cm Width: 60–70 cm Height: 100–120 cm	<20 km/h	Electric bicycle for food delivery 
ET	Length: <300 cm Width: <100 cm Height: <140 cm	<20 km/h	Electric tricycle for parcel delivery 

travel speed in accordance with the norms of electric bicycles should not exceed 20 km/h.

The development of NVM brings great convenience to people’s daily life and last mile trip. However, the rising of NMV quantity on bicycle lane leads to the increase of NMV traffic conflict which raises the accident risk and affects traffic efficiency. This paper analyzes NMV conflicts on bicycle lane using statistical analysis and regression analysis. The video logs are collected from three selected isolated bicycle lanes in Beijing. Two-dimensional trajectory tracking method is applied to characterize the NMV flow. The relationship between various traffic conflicts and traffic density as well as reverse riding ratio is examined, respectively.

2 Literature Review

Previous studies about bicycle lane safety mostly focus on the accident evaluation and conflict analysis. Bicycle accident evaluation involving traffic accident risk and the injury severity are usually conducted based on the accident data received from local traffic police departments or hospitals [11, 12]. The crash-based modeling approach can be more intuitive to reflect the traffic safety, but it needs long time collection and omission involved in the statistical period. Compared with the accident data analysis, Traffic Conflict Technique (TCT) is usually used to investigate the road traffic accident risk with the advantages in large sample and short period characteristics [13]. Therefore, conflicts are usually used as surrogates of collisions for safety analysis of a road entity, like intersection and road section [14–16].

Through the review of the previous literatures [17, 18], bicycle conflicts or accidents under mixture traffic flow receives increased attention. And the urban intersections are always selected for the conflict analysis. The influential factors of bicycle-vehicle at urban intersection were explored by an ordered logit methodology, especial the gender factor [19]. The safety of five bicycle facility layouts in intersection was compared, and it indicated that a recessed bicycle track provide a higher level of safety for cyclists [20]. However, NMV conflicts have attracted more and more attention with the development of NMVs and the construction of bicycle facility layouts.

The traffic conflicts on road segment are connected with motor vehicles, other riders or pedestrians [21, 22]. And the naturalistic cycling approach is always used to explore the bicycle conflicts, like the analysis of potential differences between conventional bicycles and electric bicycles in traffic conflicts [23]. Using the naturalistic cycling approach is convenient for obtaining conflict parameters, like the time to conflict at breaking time, conflict speed and post encroachment time, etc. [24], but its instrument is relatively complex. In addition, the video observation method is also applied to explore the traffic conflict with easier recording instrument and repeated observation [9].

This paper will analyze different NMV traffic conflicts based on TCT. The segments of urban bicycle lane in China will be selected as the observation situation. Three types of participants of the traffic conflict including bicycle, electric bicycle, and electric tricycle are considered in this paper.

3 Overall Methodology

Conflict analysis of NMV is based on the video data collection on different bicycle lanes. The process of this study can be described into following five steps as shown in Fig. 1. First, three isolated bicycle lane segments with different widths are selected. Second, the video of NMV flow is collected by placing camera on pedestrian bridge. Third, the riding track of all NMV collected in the video is traced by the trajectory tracking software. Extracted data include individual travel speed, the number of traffic conflict (TC), traffic volume, etc. Based on the data, preliminary analysis is conducted to discover the relationship among related traffic factors. Finally, correlation analysis and regression analysis are conducted to find the relationship between TC and traffic factors selected.

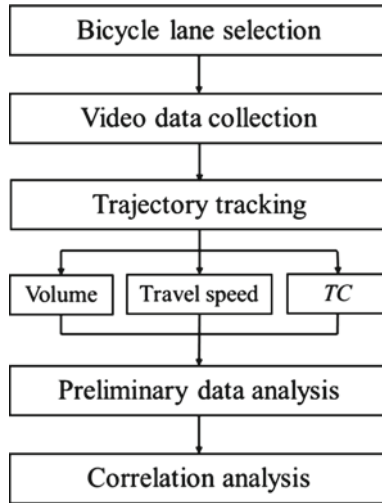


Fig. 1 Overall methodology of conflict analysis

4 Research and Data Collection of NMV Conflicts

4.1 Selecting Bicycle Lanes

The AASHTO Guild has defined the bicycle lane, “a portion of a roadway which has been designated by striping, signing, and pavement markings for the preferential or exclusive use of bicyclists” [24]. This study focuses on the traffic conflict of NMV on bicycle lane segment with different widths. In this study, three bicycle lanes whose widths are 3.4 m, 3.7 m, 4.7 m, respectively, are selected for data collection. All the selected lanes are physically isolated from motorized traffic lanes and pedestrian sidewalk. The description of the three segments is shown in Table 2 and Fig. 2.

Table 2 The description of three selected bicycle lanes

No.	Street name	Width (m)	Location
1	Chaoyangmen Street	3.4	Urban area within 2nd ring road, nearby the workplace
2	Jinsong Street	3.7	Urban area between 2nd and 3rd ring road, nearby the residential area
3	Xidawang Street	4.7	Urban area between 3rd and 4th ring road, Nearby school

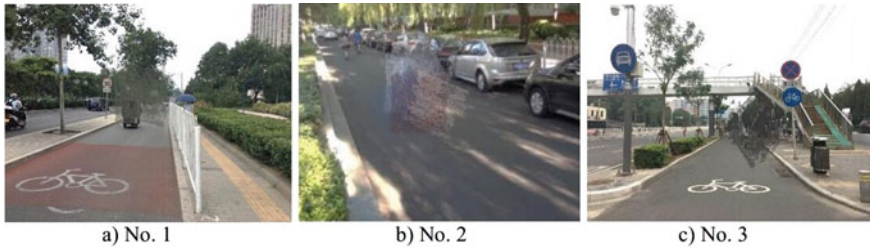


Fig. 2 Three selected bicycle lanes

4.2 Video Log Recording and Data Extraction

The camera is placed on pedestrian bridge with certain height above the bicycle lane. The camera's recording angle is adjusted to ensure the recording range including at least 50 m long bicycle lane segment. Moreover, bus stations and intersections are excluded from the bicycle lane segments selection. In this study, the video is collected in Beijing in peak hours on sunny weekdays.

NMV traffic flow performs different characteristics affected by various factors, like bicycle size, proportions of different types of bicycles [9]. In order to explore the relationship between conflict and characteristics of NMV traffic flow, traffic volume, traffic speed, and TC are collected on three selected bicycle lanes by trajectory tracking of 2820 NMVs.

5 NMV Traffic Flow Analysis

The preliminary summary and analysis of NMV traffic flow will be conducted in this section.

5.1 Traffic Volume

Traffic volumes on three bicycle lanes are 1006 veh/h, 991 veh/h and 823 veh/h, respectively. As some of NMV possess reverse riding behavior, the reverse riding ratio defined by the percentage of reverse riding NMV is recorded in Table 3.

The proportions of three types of NMV are also different, and both of the proportions of electric bicycle on Bicycle lane 1 and Bicycle lane 2 is higher than 50%, while the proportion of bicycle on Bicycle lane 3 is highest. Moreover, the percentage of electric tricycle on these three bicycle lanes is no more than 15%, as it on Bicycle lane 3 is highest with 10.21% in Fig. 3. That may be related to rider's travel activity nearby these bicycle lanes.

Table 3 Traffic volume of NMV on three bicycle lanes

Bicycle lane	Bicycle (veh/h)	Electric bicycle (veh/h)	Electric tricycle (veh/h)	Total (veh/h)	Total reverse riding ratio
Bicycle lane 1	456 (39)	513 (30)	37 (3)	1006 (72)	13.00%
Bicycle lane 2	407 (94)	503 (89)	81 (20)	991 (203)	20.48%
Bicycle lane 3	423 (69)	316 (27)	84 (11)	823 (107)	7.16%

() indicates the reverse traffic volume

The bold indicates that Bicycle lane 2 has the largest *Total reverse riding ratio*

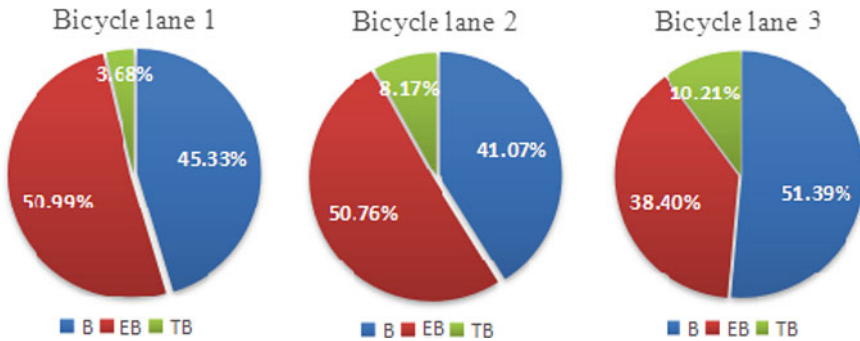


Fig. 3 Traffic volume share of NMV on three bicycle lanes

5.2 Travel Speed

The speeding phenomenon of every types of NMV is obviously observed on bicycle lane. Different types of NMV on bicycle lanes in accordance with the speed from high to low order are electric bicycle, electric tricycle, and bicycle, in Table 4.

It is found that travel speed of three types of NMV complies with the Gaussian distribution, for instance, Fig. 4 shows the distribution of electric travel speed on Bicycle lane 2, and the 85th percentile speed which is always defined to the limit speed of traffic flow is greater than the limit speed 20 km/h proposed in design specification. That also signifies that speeding characteristics no-motor vehicle flow on bicycle lanes with physical isolation, which may have a certain impact on traffic conflicts.

Table 4 Travel speed of three types of NMV on three bicycle lanes

No.	Type of NMV	Mean (km/h)	SD (km/h)	Min (km/h)	Max (km/h)	Speed limit (km/h)
1	B	20.51	4.77	8.65	37.74	15
	EB	30.40	5.59	13.00	44.06	20
	ET	25.70	6.25	12.47	39.67	20
2	B	15.84	3.11	8.43	24.04	15
	EB	26.04	5.94	7.51	41.20	20
	ET	20.13	6.33	4.65	38.57	20
3	B	16.62	4.35	10.24	37.39	15
	EB	30.05	7.82	10.57	61.66	20
	ET	26.18	6.82	12.56	46.08	20

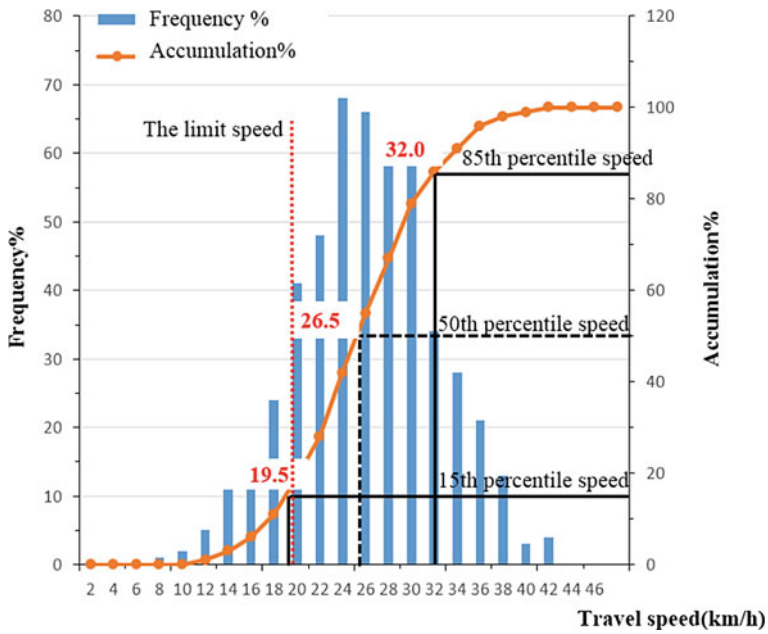


Fig. 4 Gaussian distribution of EB travel speed on Bicycle lane 2

5.3 Traffic Conflict

5.3.1 Definition and Classification of Conflicts

As the development of TCT, the traffic conflict has already been defined. “A traffic conflict is an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged” [25]. In this study, traffic conflict of NMV is redefined to determine the occurrence of a conflict in trajectory tracking process, for example in Fig. 5.

Definition: Target B and Target T conflict if both of two conditions occur:

Condition 1: $dist(B, T) \leq 0.5m$;

Condition 2: $\left| \left\{ t : \frac{dv_y}{dt} < 0 \right\} \right| \geq \frac{1}{3}t_{total}$ or

$$\left| \left\{ t : \frac{dv_y}{dt} > 0 \right\} \right| \geq \frac{1}{3}t_{total}; \left| \left\{ t : \frac{dv_x}{dt} < 0 \right\} \right| \geq \frac{1}{3}t_{total};$$

$$\left| \left\{ t : \frac{dv_x}{dt} > 0 \right\} \right| \geq \frac{1}{3}t_{total}$$

where,

$dist(B, T)$ is the distance between target B and target T contained in a traffic conflict;

v_x refers to the speed in the x direction;

v_y refers to the speed in the y direction;

t_{total} refers to the trajectory tracking time for one target;

$|\cdot|$ refers to the cardinality of a set.

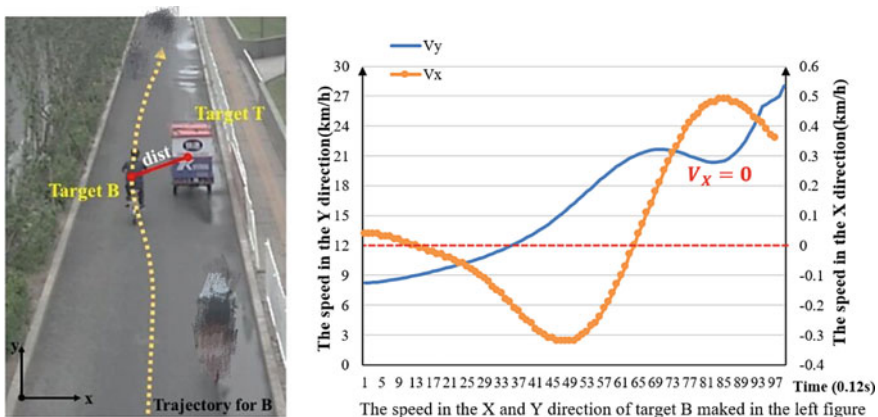


Fig. 5 An example of NMV conflict between target B and target T

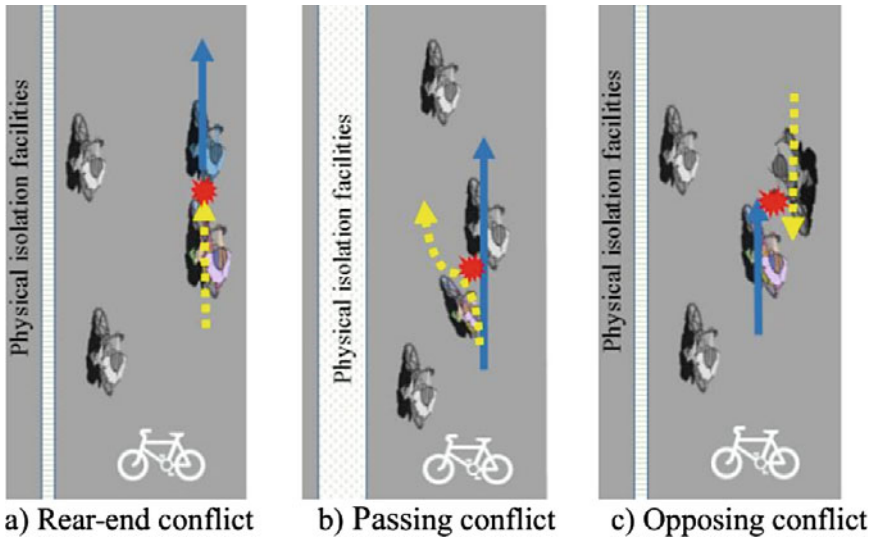


Fig. 6 Three types of the NMV traffic conflict

Traffic conflicts can be divided into different types according to different classification criteria. In this paper, traffic conflicts of NMV are classified into three types, rear-end conflict, passing conflict, and opposing conflict based on different scenes in Fig. 6. Different measures are always adopted to avoid the collision in different conflict scenes. Moreover, these three types of conflicts are various as the participants of every conflict, such as conflict between bicycle and electric bicycle.

5.3.2 Statistic Analysis of Conflicts

Table 5 shows various types of traffic conflicts. First, conflicts on three width bicycle lanes will be represented. For the total quantity of conflicts counted in 1 h among 50 m long of bicycle lane, the occurrence of traffic conflict on Bicycle lane 1 and Bicycle lane 2 is more frequently than Bicycle lane 3, with TC of 535, 487, and 212. On Bicycle lane 1, B-EB passing conflict is as high as 130 times. That indicates a potential relationship between the width of bicycle lane and the traffic volumes of both B and EB. On Bicycle lane 2, 60 times of EB-EB rear-end conflict is observed period. That is because the property of electric bicycle on this bicycle lane is higher than 50%. The opposing traffic conflict occurring on Bicycle lane 1 is more frequent than other two bicycle lanes, probably due to its higher reverse riding ratio. On Bicycle lane 3, the maximum number of three types of traffic conflict is 36, and it is the B-B rear-end conflict.

Considering the participants in conflicts, it reveals that the conflict between the non-motorized vehicles with the largest volume and others is relatively more frequent,

Table 5 All kinds of TC on three bicycle lanes

No.	Conflict type	Passing	Rear-end	Opposing	Total
1	B-B conflict	26	49	15	90
	B-E conflict	130	19	52	201
	B-T conflict	8	5	1	14
	E-E conflict	33	47	25	105
	E-T conflict	7	6	9	22
	T-T conflict	0	1	3	4
	Total	204	127	105	436
2	B-B conflict	22	51	11	84
	B-E conflict	82	26	50	158
	B-T conflict	12	18	18	48
	E-E conflict	36	60	43	139
	E-T conflict	16	15	19	50
	T-T conflict	1	2	5	8
	Total	169	172	146	487
3	B-B conflict	20	36	6	62
	B-E conflict	29	7	15	51
	B-T conflict	9	12	7	28
	E-E conflict	14	17	5	36
	E-T conflict	9	8	8	25
	T-T conflict	4	2	4	10
	Total	85	82	45	212

especially on Bicycle lane 1 and Bicycle lane 2. However the frequency of conflict between ET and others is not obvious compared with other types of conflicts. This is mainly due to ET occupied the lower travel sharing rate and bigger road space. The rear-end conflict occurs more frequent than other two kinds, passing conflict and opposing traffic between the same kind of NMV in Fig. 7.

In this portion, the parameters of NMV on bicycle lanes, like traffic volume, travel speed, and the number of conflict have been investigated. The property of different types of NMV on three width bicycle lanes is various, as bicycle and electric bicycle are the mainly participant on these three observed bicycle lanes. Furthermore, the reverse riding phenomenon on Bicycle lane 2 is more obvious than other two bicycle lanes. The travel speed of NMV on bicycle lanes follows the Gaussian distribution, and the speeding phenomenon is very common. The three types of traffic conflict between different participants occur with some regularity. It is necessary to explore the specific relationship between conflict and other parameters of NMV.

In this section the correlation relationship of TC and these considered factors of NMV flow are explored by Spearman rank correlation analysis [26]. The Spearman rank correlation analysis is regardless of the overall distribution of the two variables

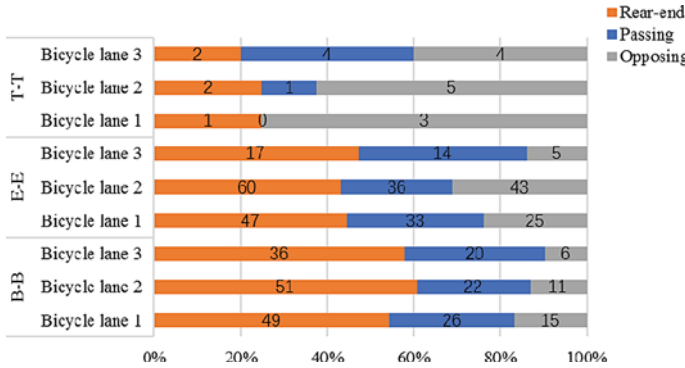


Fig. 7 Percentage analysis of conflicts between the same types of NMV

and the size of the sample, compared with other correlation analysis methods, and it is suited to study by regarding TC and other characteristics of NMV flow as ordinal variables. The remarkable relative factors (which are marked bold in Tables 5, 6, 7, 8 and 9) are selected with the value of the significance test less than 0.05. According to experience [27], the correlation degrees are divided into the following four grades.

Based on the data analysis in the previous section, these main characteristic factors of NMV flow have been extracted including traffic volume, travel speed, and reverse riding ratio of different types of NMV. Travel speed and traffic volume usually are connected by the traffic density according to traffic flow theory in Eq. 1 [28]. The correlation analysis TC is correlatively analyzed with traffic density of different NMVs and reverse riding ratio of different NMV respectively.

$$Q = KV \tag{1}$$

Table 6 Correlation analysis of traffic density and TC

Conflict	K_B		K_{EB}		K_{ET}	
	ρ	Sig.	ρ	Sig.	ρ	Sig.
Passing	0.151	0.380	0.661**	0.000	-0.192	0.262
Rear-end	0.346*	0.039	0.718**	0.000	0.082	0.634
Opposing	0.292	0.083	0.676**	0.000	0.034	0.846

Uncorrelated: $|\rho| \leq 0.30$

Weak correlation: $0.30 < |\rho| \leq 0.65$, marked with *

Moderate correlation: $0.65 < |\rho| \leq 0.80$, marked with **

Strong correlation: $0.80 < |\rho| \leq 1.00$, marked with ***

where,

K_i refers to the traffic density of NMV i;

ρ refers to the spearman rank correlation coefficient;

Sig. refers to the value of the significance test.

Table 7 Correlation analysis of traffic density and the TC

Conflict	K_B		K_{EB}		K_{ET}	
	ρ	Sig.	ρ	ρ	Sig.	ρ
B-B	0.645*	0.000	0.119	0.491	-0.225	0.188
B-EB	0.247	0.146	0.574*	0.000	-0.236	0.166
B-ET	0.289	0.088	0.135	0.434	0.512*	0.001
EB-EB	-0.025	0.886	0.896***	0.000	-0.164	0.338
EB-ET	0.112	0.515	0.361*	0.031	0.573*	0.000
ET-ET	-0.111	0.520	-0.124	0.471	0.591*	0.000

Table 8 Correlation Analysis of Reverse Riding Ratio and TC

Conflict	R_B		R_{EB}		R_{ET}	
	ρ	Sig.	ρ	Sig.	ρ	Sig.
Passing	-0.042	0.846	0.042	0.846	-0.100	0.641
Rear-end	0.210	0.326	0.111	0.606	0.469	0.021
Opposing	0.775**	0.000	0.573*	0.003	0.511*	0.011

where,

R_i refers to the reverse riding ratio of NMV i.

Table 9 Correlation analysis of reverse riding ratio and TC

Conflict	R_B		R_{EB}		R_{ET}	
	ρ	Sig.	ρ	Sig.	ρ	Sig.
B-B	0.077	0.656	-0.05	0.774	-0.237	0.164
B-EB	-0.155	0.366	0.123	0.474	-0.194	0.257
B-ET	0.511*	0.01	0.405*	0.014	s	0.038
EB-EB	0.098	0.568	0.451*	0.006	0.103	0.55
EB-ET	-0.028	0.873	0.265	0.119	0.428*	0.009
ET-ET	-0.017	0.922	0.126	0.463	0.345*	0.039

where,

Q refers to the traffic volume;

K refers to the traffic density;

V Refers to the Travel Speed.

5.4 Correlation Analysis

5.4.1 TC and Traffic Density

NMV conflicts are classified into passing conflict, rear-end conflict and opposing conflict based on different forms, and they are various with different participates in the mythology section.

First, TC in these three types of NMV conflict is correlatively analyzed with traffic density of B, EB, and ET respectively. The results show that all of these three conflicts have moderate correlation with EB density in Table 6. In addition, rear-end confliction has weak correlation with B density.

Second, six types of NMV conflict with different NMV participates, including B-B conflict, B-EB conflict, B-ET conflict, EB-EB conflict, EB-ET conflict, and ET-ET conflict, are considered to explore the correlation with B density, EB density, and ET density. It indicates that different types of NMV conflict are correlative with the density of different types of NMV in Table 7.

EB-EB conflict has strong correlation with EB density, and they are fitted into a linear function in Eq. 2 with the coefficient of determination of 0.769. It could be considered to predict the EB-EB conflict on EB mainly involved bicycle lane.

$$TC_{EB-EB} = 10.395K_{EB} - 5.64 \quad (2)$$

where,

TC_{EB-EB} is EB-EB conflict.

5.4.2 TC and Reverse Riding Ratio

The correlation analysis of three types of NMV conflict and reverse riding ratio are conducted in Table 8. The results show that opposing conflict has weak correlation with reverse riding ratio. In order to further explore the relationship between reverse riding ratio and different conflicts with various NMV participants, the correlation analysis is conducted.

$$TC_{op} = 7.836 + 91.781R_B + 36.219R_{EB} + 40.611R_{ET} \quad (3)$$

where,

TC_{op} refers to TC of opposing conflict.

The regressing analysis is conducted based on the correlation analysis, between TC_{op} and R_B , R_{EB} , R_{ET} in Formula 2. The coefficient of determination is 0.704 which shows a better fitting result. The multivariate function could be applied to predict the opposing conflict according to the reverse riding ratio of different NMVs in real world.

The correlation analysis of NMV conflict occurred between different types of NMV, and reverse riding ratio are conducted in Table 9. But the results show that there is weak correlation between reverse riding ratio and B-B conflict, B-EB conflict, and other types of NMV conflict have general correlation with reverse riding ratio. This indicates that the occurrence of traffic conflicts involving NMV with large width has a certain relationship with reverse riding ratio. This is because they take up more road space.

6 Conclusions and Recommendations

Recent years, the riding environment of bicycle lanes is gradually paid more attention with the rapid development of various NMVs in China. This paper mainly explores the traffic conflict of various NMVs on bicycle lanes to recommend more friendly and safe riding environment. The relationship of traffic conflict and characteristics of NMV traffic flow is investigated by correlation analysis and regression analysis.

The increasing of traffic density on bicycle lane will lead to the increase of corresponding conflict according to the correlation coefficient. The obvious speeding riding brings up potential risk for bicycle lane users. The reverse riding ratio of NMV also have great influence for the traffic conflict especially for the opposing conflict. It recommends that speeding and reverse riding of NMVs should be restricted by NMV laws or regulations to improve the safety and traffic efficiency of bicycle lane. And non-motorized lanes with two-way traffic should be replaced by unidirectional non-motorized lanes with physical isolation, especially in urban areas with large non-motorized traffic flow.

The findings of this study reflect the influence of NMV flow characteristic factors in traffic conflict on bicycle lane segment. However, the limitations in this study are that only three types of bicycle lines were observed, and other interference factors of NMV flow were idealized. In further exploration, more types of bicycle lanes with various width and interference situation will be taken into consideration. Furthermore, the relationship of electric bicycle and more potential factors will be explored, like specific electric bicycle performance and rider behaviour characteristics.

Acknowledgements This research was funded by the Key Research Plan of Shandong Province China (2019GSF109057).

References

1. iiMedia Research (2018) China's shared bicycle development status research. *J Environ Econ Z2*(2017):78–79
2. Muetze A, Tan YC (2007) Electric bicycles—a performance evaluation. *IEEE Ind Appl Mag* 13(4):12–21

3. Zhao J, Liu M, Liu L, Wang S, Zhang Z (2014) Safety test of electric bicycle under human-car-road system. *China J. Highway Transp.* **27**(2):105–111
4. Feng ZY, Raghuwanshi RP, Xu ZG, Huang DY, Chong Z, Tao J (2010) Electric-bicycle-related injury: a rising traffic injury burden in China. *Injury Prevent J Int Soc Child Adolesc Injury Prevent* **16**(6):417
5. *Urban Road Design Code (CJJ37-95)* (1991) China Construction Industry Press
6. Johnson M, Rose G (2013) Electric bikes—cycling in the New World City: An investigation of Australian electric bicycle owners and the decision making process for purchase. *Australasian Transport Research Forum (ATRF)*, 36th, 2013, Brisbane, Queensland, Australia
7. Rose G (2012) E-bikes and urban transportation: emerging issues and unresolved questions. *Transportation* **39**(1):81–96
8. Toole JL, Zimny B (2012) Bicycle and pedestrian facilities. *Cur Org* (2012)
9. Zacharias J, Zhang B (2015) Local distribution and collection for environmental and social sustainability—tricycles in central Beijing. *J Transp Geogr* **49**:9–15
10. Kröyer HRG (2015) The relation between speed environments, age and injury outcome for bicyclists struck by a motorized vehicle—a comparison with pedestrians. *Accid Anal Prev* **76**:57–63
11. Eluru N, Bhat CR, Hensher DA (2008) A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accid Anal Prev* **40**(3):1033–1054
12. Kočárková D (2012) Traffic conflict techniques in Czech Republic. *Procedia Soc Behav Sci* **53**:1028–1033
13. Perkins SR, Harris JI (1968) Traffic conflict characteristics: accident potential at intersections. *Psychol Rep* **35**(1):73–74
14. John S, Spicer BR (1976) Traffic conflicts—a development in accident research. *Hum Fact J Hum Fact Ergon Soc* **18**(4):335–350
15. Williams MJ (1981) Validity of the traffic conflicts technique. *Accid Anal Prev* **13**(2):133–145
16. Jutaek O, Eungcheol K, Myungseob K, Sangho C (2010) Development of conflict techniques for left-turn and cross-traffic at protected left-turn signalized intersections. *Saf Sci* **48**(4):460–468
17. Autey J, Sayed T, Zaki MH (2012) Safety evaluation of right-turn smart channels using automated traffic conflict analysis. *Accid Anal Prev* **45**(1):120–130
18. Stipančić J, Zangenehpour S, Miranda-Moreno L, Saunier N, Granié MA (2016) Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology. *Accid Anal Prev* **97**:19–27
19. Madsen TKO, Lahrman H (2016) Comparison of five bicycle facility designs in signalized intersections using traffic conflict studies. *Transp Res Part F Psychol Behav*
20. Chen J, Xie ZQ (2009) Cycle traffic conflict model on urban pedestrian-bicycle paths. *J Jilin Univ* **39**:121–125
21. Masaru K, Upali V (2000) Bicycle and pedestrian traffic conflicts on shared pavements
22. Petzoldt T, Schleinitz K, Heilmann S, Gehlert T (2016) Traffic conflicts and their contextual factors when riding conventional vs. electric bicycles. *Transp Res Part F Psychol Behav*
23. Angel-Domenech A, Garcia A, Agustin-Gomez F, Llorca C (2014) Traffic conflict analysis using an instrumented bicycle on cycle tracks of Valencia. In: *International Cycling Safety Conference*
24. Wang D, Zhou D, Jin S, Ma D (2015) Characteristics of mixed bicycle traffic flow on the conventional bicycle path. *Transportation Research Board 94th Annual Meeting*
25. Doran JW, Linn DM (1979) *Proceedings of the first workshop on traffic conflicts: (held September 1977 in Oslo.) institute of transport economics, Oslo, Norway, 1977. 138 pp.* *Accident Anal Prevent* **11**(4):317–318
26. Zar JH (2004) Spearman rank correlation
27. Shao C (2016) *Traffic survey and data analysis.* China Communications Press Co., Ltd.
28. Drew D (1968) *Traffic flow theory and control.* Highways

Parameter Matching and Simulation Analysis on Power System of Wheel-Legged Vehicle



Xizheng Chen, Xiaobei Jiang, Lijin Han, Wenlin Yu, Jiawen Guo, and Wuhong Wang

Abstract With the development of multi-body dynamics and robotics, wheel-legged vehicles have raised attentions in the field of special vehicles for their high flexibility and adaptability. Based on the existing model of wheel-legged vehicle power system, a kinetics analysis model of the power system is proposed in this study, and the model parameters are matched and analyzed under different working conditions. Vehicle simulation model is designed by Advisor 2002 software, and the dynamic performance of the vehicle is analyzed under the ECE-EUDC cycle conditions. The results show the adaptability of wheel-footed vehicles to different road conditions. The practicability and feasibility of the wheel-footed vehicle program is proved.

Keywords Wheel-legged vehicle · Power system · Parameter matching · Simulation

1 Introduction

As the most used means of transportation and delivery by humans, vehicles play an important role in social life. However, limited by its single function, traditional vehicles can no longer meet the increasing demands of mankind. The development of utility vehicle with high adaptability and flexibility has become one of the main tasks of present car companies. Wheel-legged vehicles combine the excellent performance of robots and traditional vehicles. It is an innovation and a landmark concept [1].

Space detection robot is the main research direction of wheel-legged vehicles in the early days. The “Sojna” Mars rover developed and manufactured by the US Jet Propulsion Laboratory in the 1990s is one of the early representative products. It combines manipulator technology with wheels to effectively improve the vehicle’s mobility. Hyundai Motor Company launched an innovative concept car called “Elevate” at the 2019 CES show. The top of the four foldable robotic arms of the car is equipped with wheels, which can drive, so that the car can walk or climb when facing

X. Chen · X. Jiang (✉) · L. Han · W. Yu · J. Guo · W. Wang
School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China
e-mail: jiangxiaobei@bit.edu.cn

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_21

291

difficult terrain. Hyundai Motor stated that this vehicle, which combines electric vehicle technology and robotics technology, is the ultimate mobile car.

The dynamic performance of wheel-legged vehicles is as important as traditional vehicles. The performance of the power system directly determines the performance of the vehicle. Therefore, it is of great significance to design a properly matched power system. In this paper, the dynamic analysis of the power system of the wheel-legged vehicles is carried out, and the parameters of the in-wheel motor and battery pack are determined according to different working conditions. The dynamic analysis of the vehicle model based on Advisor is carried out, and the simulation results show that the design of the power system is reasonable. Therefore, the design and development of wheel-legged vehicles has high practical application value [2].

2 Power System Design

The power system is one of the most important parts of the vehicle, which directly determines the performance of the vehicle. Different from the power transmission system of traditional vehicles, the power system of wheel-legged vehicles is more similar to that of wheeled robots. The power system of wheel-legged vehicles is directly powered by the wheel hub motor. The determination of power system parameters includes preliminary design and performance check [3].

2.1 Wheel-Legged Vehicle Dynamics Analysis

Wheel-legged vehicles will have resistance that affects their driving during normal driving, mainly including rolling resistance F_f , air resistance F_w , and gradient resistance F_i . The total resistance of a wheel-legged vehicle under normal driving conditions is:

$$\sum F = F_f + F_w + F_i \quad (1)$$

$$\begin{cases} F_f = fmg \\ F_w = \frac{C_D A u_a^2}{21.15} \\ F_i = G \sin \alpha \end{cases} \quad (2)$$

- f —Rolling resistance coefficient;
- m —The mass wheeled vehicles;
- g —Gravitational acceleration;
- C_D —Drag coefficient;
- A —Frontal area;
- u_a —Velocity;

α —Slope angle.

In different working conditions, the driving force F_t required by the vehicle is different, which can be derived from the vehicle driving equation and the total resistance equation of the vehicle traveling in a straight line:

$$m \frac{du}{dt} = F_t - (F_f + F_w + F_i) \quad (3)$$

When the total resistance is equal to the driving force of the vehicle, the driving force and the resistance are offset, and the wheel-legged vehicle will drive at a constant speed. When the total resistance is less than the driving force of the vehicle, the remaining part of the driving force after offsetting the resistance will accelerate the vehicle. When the total resistance is greater than the driving force, the vehicle will slow down.

Therefore, the driving conditions that satisfy the normal driving of a wheel-legged vehicle are:

$$F_t \geq (F_f + F_w + F_i) \quad (4)$$

Driving conditions are only necessary conditions for vehicle driving, but they are not sufficient. Adhesion conditions are also important. If the vehicle cannot reach the driving state, the dynamic performance will be limited.

The attachment conditions of the vehicle are:

$$F_t \leq \varphi \cdot F_Z \quad (5)$$

F_Z —Vertical load;

φ —Vertical load factor.

Taken together, the necessary and sufficient conditions for vehicle driving can be obtained:

$$(F_f + F_w + F_i) \leq F_t \leq \varphi \cdot F_z \quad (6)$$

2.2 The Power Plant

2.2.1 Power Analysis of Driving Motor

The hub motor is the main component of the wheel-legged vehicle power system. The greater the power that the motor can provide, the greater the dynamic performance of the vehicle, including the maximum speed, off-road performance and climbing ability. Therefore, the peak power of the motor must meet the requirements of the maximum driving speed power P_{max1} , the maximum grade power P_{max2} and the

maximum off-road driving speed power P_{max3} . Therefore, the peak power P_{max} is:

$$P_{max} \geq \max(P_{max1}, P_{max2}, P_{max3}) \quad (7)$$

$$P = \frac{v_i}{3600\eta_t} \left(mgf \cos \alpha_{max} + mg \sin \alpha_{max} + \frac{C_D A v_i^2}{21.15} \right) \quad (8)$$

η_t —Electrical efficiency.

According to the characteristic curve of the motor, the motor needs to go through the constant torque section and constant power section during operation. When the vehicle is running, constant torque is output at low speed, and constant power is output at high speed. Therefore, the rated power is the power corresponding to the maximum vehicle speed.

2.2.2 Speed Analysis of Driving Motor

Hub motor directly drives wheel-legged vehicles, therefore, the rated speed corresponds to the common driving speed of the vehicle, and the highest speed needs to meet the requirements of the highest driving speed of the vehicle.

40–60 km/h is the most commonly driving speed for vehicles in urban conditions, and the highest driving speed is determined according to the design requirements. Calculate the motor speed according to the following formula:

$$n = \frac{1000v}{60\pi D} \quad (9)$$

D —Wheel diameter.

2.2.3 Analysis of Battery Pack

The battery is the only energy component of the wheel-legged vehicle and the core component of the vehicle, which determines the performance of the vehicle in many aspects. Lithium iron phosphate battery technology is mature, with good security and charge-discharge efficiency. It can meet the needs of wheel-legged vehicle.

The battery energy index is one of the important parameters reflecting the value of the battery. The energy of the battery pack determines the cruising range of the wheel-legged vehicle. The greater the energy of the battery pack, the longer the cruising range and the longer the working time.

The required power P for wheel-legged vehicles can be calculated by the following formula:

$$P = \frac{v}{3600\eta_t} mgf + \frac{C_D A v^2}{21.15} + P_e \quad (10)$$

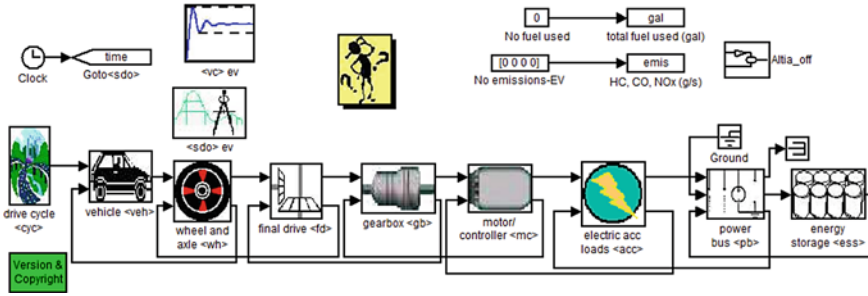


Fig. 1 Wheel-legged vehicle simulation model

P_e represents the power required by the wheeled vehicle to carry other electrical equipment.

3 Simulation Model Based on Advisor 2002 Software

3.1 Wheel-Legged Vehicle Simulation Model

Advisor 2020 software comes with a variety of electric vehicle models [4]. Referring to the power system layout of wheel-legged vehicles, a simulation model of the entire vehicle is built based on the EV-LG-in basic model in this simulation. The models mainly include: working condition input module, body model, wheel and axle model, drive motor module, vehicle electronic accessory module and power battery module, etc. The Wheel-legged vehicle simulation model is shown in Fig. 1.

3.2 Battery Simulation Model

When building a battery simulation model, it is necessary to calculate the open circuit voltage and internal resistance of the battery pack according to the current SOC value, battery pack temperature and required output power [5]. In order to prevent the battery current power from exceeding the limit range, it is also necessary to calculate the power limit module through the SOC value and equivalent circuit parameters. The battery model can also calculate the remaining battery capacity. Combining these functional modules can form a battery assembly model. The battery simulation model is shown in Fig. 2.

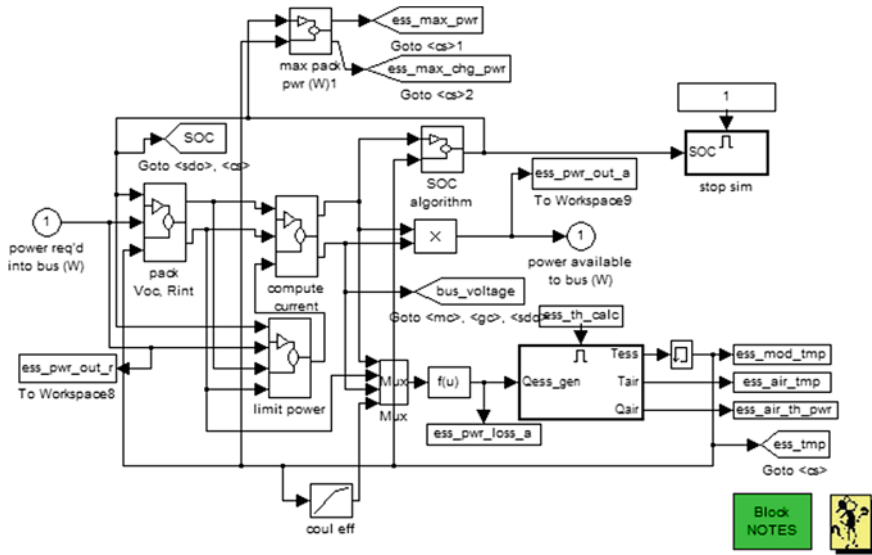


Fig. 2 Battery simulation model

3.3 Body Simulation Model

In the simulation process, various resistances are integrated on the body simulation model, which has a resistance calculation module and a driving speed calculation module. The driving force of the car can be obtained, and then the final speed of the vehicle can be obtained according to the characteristics of the motor. The body simulation model is shown in Fig. 3.

3.4 Adjustment of Simulation Model

In the simulation model of Advisor, the original parameters do not match those of wheel-legged vehicles, so it is necessary to adjust the parameters to meet the simulation requirements. Wheeled-footed vehicles are directly driven by the hub motor, and do not need the main reducer and gearbox module, so it is necessary to modify the main reduction ratio to 1 during simulation. When setting the relevant parameters of the motor, it is necessary to modify the program according to the characteristics of the motor to obtain a reasonable parameter set. It also includes wind resistance coefficient, windward area, battery parameters and tire size, etc. The simulation can be completed after inputting environmental parameters for different working conditions.

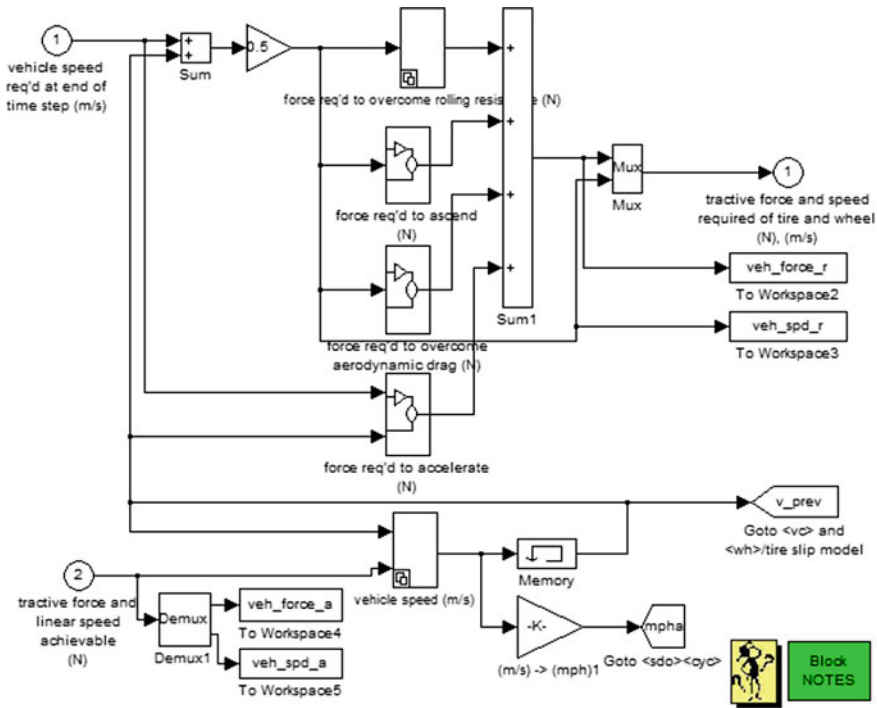


Fig. 3 Body simulation model

4 Simulation Results

Working condition method is a common method to evaluate vehicle performance, which is widely used in many countries. According to different situations, the simulation cycle conditions selected by different countries are different, and China mainly adopts ECE-EUDC cycle conditions. The ECE-EUDC cycle can be divided into two parts. The first part is the traditional urban road driving condition, which is a simplified representative of the urban driving process, and the second part is an additional suburban driving condition, which represents the suburban vehicle running process. The combination of the two working conditions can represent most of the vehicle driving conditions. Parameters of ECE-EUDC cycle are shown in the following Table 1.

The vehicle speed changes with time as shown in Fig. 4. The vehicle can run normally at different speeds such as idling, accelerating, constant speed, and decelerating, and is highly adaptable to urban road conditions and suburban road conditions. In the starting phase of the vehicle, the vehicle speed increases relatively slowly; after the vehicle is driving stably, the vehicle speed increases relatively rapidly, which is in line with the formal law of electric drive vehicles.

Table 1 ECE-EUDC parameter table

Projects	Numerical
Mileage	10.93 km
Duration	1225 s
Maximum speed	120 km/h
Average speed	32.1 km/h
Number of stops	13

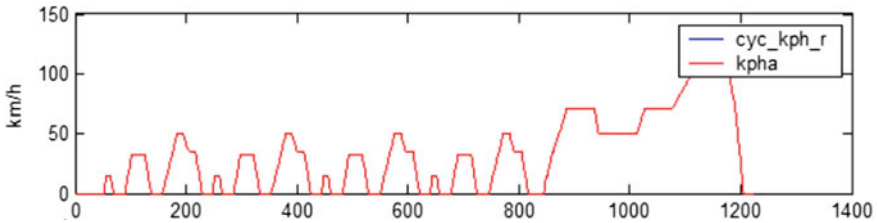


Fig. 4 The change in velocity with time

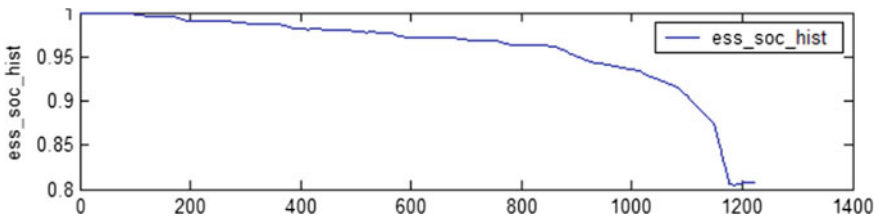


Fig. 5 The state of charge of the battery

Figure 5 is a graph of the state of charge of the battery. Under ECE-EUDC cycle conditions, the wheel-legged vehicle needs to accelerate and decelerate frequently. During the entire driving process, the state of charge of the battery decreases slowly for a long time, and the vehicle can reclaim energy to charge the battery. During the whole process, the battery has been kept in a stable working area, which can effectively guarantee the working life of the battery.

5 Summary

In this paper, the wheel-legged vehicle is taken as the research object, and its power system is designed and analyzed. Finally, the whole vehicle’s power performance is analyzed by Advisor 2002 software. The results show that the wheel-legged vehicle can complete idling, decelerating, constant-speed and accelerated driving under the ECE-EUDC cycle conditions, indicating that the wheel-legged vehicle can adapt to

urban road driving conditions and suburban driving conditions. During the operation of the wheel-legged vehicle, the battery can maintain a good charge and discharge state, which effectively ensures the normal driving of the vehicle. On the whole, the research and development of wheel-legged vehicles has high practical significance.

Further work will improve the detailed parameters of the power system, and combine the three-dimensional model for simulation to provide theoretical support for the research of wheel-legged vehicles.

Acknowledgements This research was partially funded by the National Project of China (2019-JCJQ-ZD-133-01).

References

1. Ding X, Zheng Y, Xu K (2017) Wheel-legged hexapod robots: a multifunctional mobile manipulating platform. *Chinese J Mech Eng*
2. Folkesson A, Andersson C, Alvfors P, Alakuela M, Overgaard L (2003) Real life testing of a hybrid PEM fuel cell bus. *J Power Sour* 118(1/2):349–357
3. Fellini R, Michelena N, Papalambros P, Sasena M (1999) Optimal design of automotive hybrid powertrain systems. *IEEE Xplore*
4. Markel T, Brooker A, Hendricks T, Johnson V, Kelly K, Kramer B et al (2002) Advisor: a systems analysis tool for advanced vehicle modeling. *J Power Sour* 110(2):255–266
5. Boukehili A, Zhang YT, Zhao Q, Ni CQ, Su HF, Huang GJ (2012) Hybrid vehicle power management modeling and refinement. *Int J Automot Technol* 13(6):987–998

An Online Impedance Algorithm for All-Climate Batteries Used on Electric Vehicle



Tao Zhang, Xiaoxia Sun, Jun Tian, and Ning Zhao

Abstract To improve the performance of power batteries for all climate application, a novel online model-based resistance algorithm (OMRA) has been developed to obtain the impedance of batteries at different temperatures and current amplitudes. The Butler-Volmer (BV) equation has been proposed to describe the electrochemical reaction process of the battery, and then a novel BV circuit model has been employed for accurately calculating the resistance of the battery at different currents and temperatures. The theoretical results show that contrast to the traditional method, using OMRA can accurately calculate the battery resistance and maintain the maximum calculation error at less than 2 m Ω . This proposed method which requires low computational capability and storage space has the potential to calculate the impedance online and accurately.

Keywords All-climate electric vehicles · Low temperature · Online impedance algorithm · Model-based

1 Introduction

Electric vehicle technology plays a profound impact on solving the problems of energy shortage and environmental pollution [1]. The power battery, which plays a critical role in electric vehicles, is also the significant bottleneck of electric vehicles [2]. It is highly sensitive to many factors. Especially, the variability of the resistant is explicit at different temperatures and currents.

T. Zhang · X. Sun (✉) · J. Tian · N. Zhao
China North Vehicle Research Institute, Beijing 100072, China
e-mail: sun_xiaoxia1983@163.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_22

301

1.1 Literature Review

Conventionally, the electrochemical impedance spectroscopy (EIS) measurement and DC internal resistance test are used to character the battery resistance [3–8]. In [6] an AC impedance of electrochemical batteries is measured online by controlling the power converter and duty-cycle perturbation. An online impedance measurement system is proposed to measure impedance at a range of frequencies using the existing power electronics [7]. Impedance measurement is adopted to detect and analyze faults in battery packs [8]. The online method presented in [9] requires no signal injection and complicated circuits. However, the measurement of impedance spectrum which includes different frequencies needs to be done sequentially [10]. Although the two traditional methods have been widely employed in the laboratory, the equipment are relatively costly and bulky to gain available results [11]. Meantime, the internal resistance of the battery cannot be measured in real time, and many efficient models and algorithms are based on the offline data, meantime, the offline data are usually measured by the DC internal resistance test and AC impedance test which result are susceptible to the location, the length of the test instrument, the test idioms and so on.

1.2 Contributions of the Work

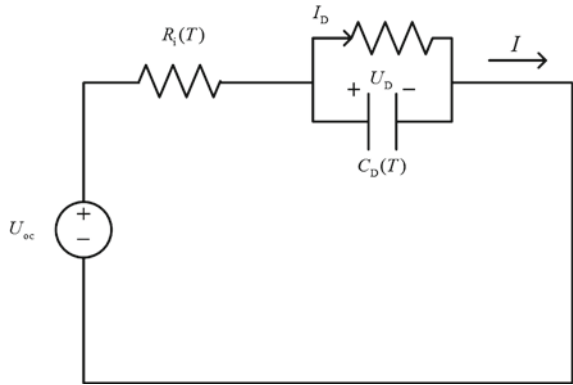
To reduce the dependence on the offline measurement equipment when obtain the battery resistance, a novel online model-based resistance algorithm has been developed to calculate the impedance of batteries. (i) The OMRA facilitates to obtain the resistance at different temperatures and currents. (ii) The proposed method calculate the resistance online without no equipment. (iii) The presented method can calculate the battery resistance continuously and accurately and maintain the maximum error within 2 mΩ.

2 Description of the Online Model-Based Impedance Method

According to the circuit model based on the B-V equation shown in Fig. 1, the polarization current I_D can be written as [12]:

$$I_D = S \cdot i_0 \left\{ \exp\left(\frac{\alpha_a \cdot F \cdot \eta}{R_g \cdot T}\right) - \exp\left(-\frac{\alpha_c \cdot F \cdot \eta}{R_g \cdot T}\right) \right\} \quad (1)$$

Fig. 1 Butler-volmer circuit model



where S is the active surface of the electrode, α_a and α_c are transfer coefficient of anode and cathode, $\alpha_a = \alpha_c = 0.5$, i_0 is exchange current density, R_g is the gas constant and η is the over-potential of the intercalation reaction, T is the temperature.

Equation (1) can be transformed as following equation by the change of trigonometric function $\sinh x = \frac{1}{2}[\exp(x) - \exp(-x)]$.

$$I_D = 2 \cdot S \cdot i_0 \cdot \sinh\left(\frac{\alpha_a \cdot F}{R_g \cdot T} \cdot \eta\right) \tag{2}$$

According to the hyperbolic sin function transformation [13] $\sinh^{-1} x = \ln[x + (x^2 + 1)^{0.5}]$, Eq. (2) can be rewritten as:

$$\eta = \ln \left\{ \frac{1}{2 \cdot S \cdot i_0} \cdot I_D + \left[\left(\frac{1}{2 \cdot S \cdot i_0} \cdot I_D \right)^2 + 1 \right]^{1/2} \right\} \cdot \frac{R_g \cdot T}{\alpha_a \cdot F} \tag{3}$$

The relationship between the polarization resistance R_D and over-potential η can be described as:

$$\eta = I_D \cdot R_D \tag{4}$$

Equations (3) and (4) can be combined as:

$$\frac{\alpha_a \cdot F}{R_g \cdot T} \cdot I_D \cdot R_D = \ln \left\{ \frac{1}{2 \cdot S \cdot i_0} \cdot I_D + \left[\left(\frac{1}{2 \cdot S \cdot i_0} \cdot I_D \right)^2 + 1 \right]^{1/2} \right\} \tag{5}$$

So, R_D can be described as:

$$R_D(I) = \frac{\ln \left\{ \frac{1}{2 \cdot S_{i0}} \cdot I_D + \left[\left(\frac{1}{2 \cdot S_{i0}} \cdot I_D \right)^2 + 1 \right]^{1/2} \right\}}{\frac{\alpha_a \cdot F}{R_g \cdot T} \cdot I_D} \quad (6)$$

According to Arrhenius equation [14] $R_D(T) = A \exp\left(\frac{E_a}{T}\right)$, the R_D can be rewritten as:

$$R_D(T, I) = A \cdot \exp\left(\frac{E_a}{T}\right) \cdot \frac{\ln \left\{ \frac{1}{2 \cdot S_{i0}} \cdot I_D + \left[\left(\frac{1}{2 \cdot S_{i0}} \cdot I_D \right)^2 + 1 \right]^{1/2} \right\}}{\frac{\alpha_a \cdot F}{R_g \cdot T} \cdot I_D} \quad (7)$$

where E_a is the activation energy, A is a constant.

The total resistance can be depicted as:

$$R(T, I, f) = R_i(T) + \frac{A \cdot \exp\left(\frac{E_a}{T}\right) \cdot \frac{\ln \left\{ \frac{1}{2 \cdot S_{i0}} \cdot I_D + \left[\left(\frac{1}{2 \cdot S_{i0}} \cdot I_D \right)^2 + 1 \right]^{1/2} \right\}}{\frac{\alpha_a \cdot F}{R_g \cdot T} \cdot I_D}}{1 + (2\pi f)^2 \left(A \cdot \exp\left(\frac{E_a}{T}\right) \cdot \frac{\ln \left\{ \frac{1}{2 \cdot S_{i0}} \cdot I_D + \left[\left(\frac{1}{2 \cdot S_{i0}} \cdot I_D \right)^2 + 1 \right]^{1/2} \right\}}{\frac{\alpha_a \cdot F}{R_g \cdot T} \cdot I_D} \right)^2} C_D^2 \quad (8)$$

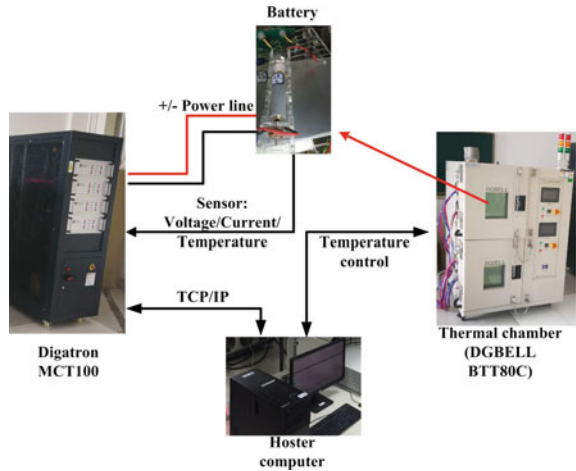
In the light of the above formula derivation, the battery resistance is affected by many factors such as temperature, current, frequency and so on.

3 The Experiment

3.1 The Experimental Set Up

All test cells were the commercial 2.4 Ah NMC lithium-ion batteries. All tests were carried out by a battery test system (Digatron MCT-100). The environmental temperature was controlled by a thermal chamber (DGBELL-BTT80C), human-machine interaction and experimental data collections were accomplished by a host computer. All is shown in Fig. 2.

Fig. 2 The configuration of the experiment



3.2 Measurement of the Hybrid Pulse Power Characteristic (HPPC)

To investigate the thermo-sensitivity and amplitude-sensitivity of battery resistance, the test batteries with 50% SOC (state of charge) are soaked in the chamber set and fixed at 0/10/20/30 °C respectively for more than four hours. The HPPC is a sequence of the pulse cycles consisting of seven different charge–discharge currents, considering that the typical current ranges of the battery in EVs are between 0 and 3C [15]. The sampling interval in the experiments is 1 s. Figure 3 shows the experimental data of the current versus time for the cell under the HPPC test.

4 Online Model-Based Impedance Estimation and Discussion

An online model-based resistance algorithm of batteries is conducted at different currents and temperatures to obtain the resistance online. In this section, the resistance of the traditional method is identified by genetic algorithm based on the experimental data which is accepted widely. The OMRA results are calculated by Eq. (8). The results of the two methods of the three batteries with different SOC are presented and discussed. Figure 4 shows battery resistances changes with the variability of currents and temperatures. The reliability and validity of the proposed method are verified by the mathematical model and the traditional method. And the maximum error of OMRA is lower than 2 mΩ.

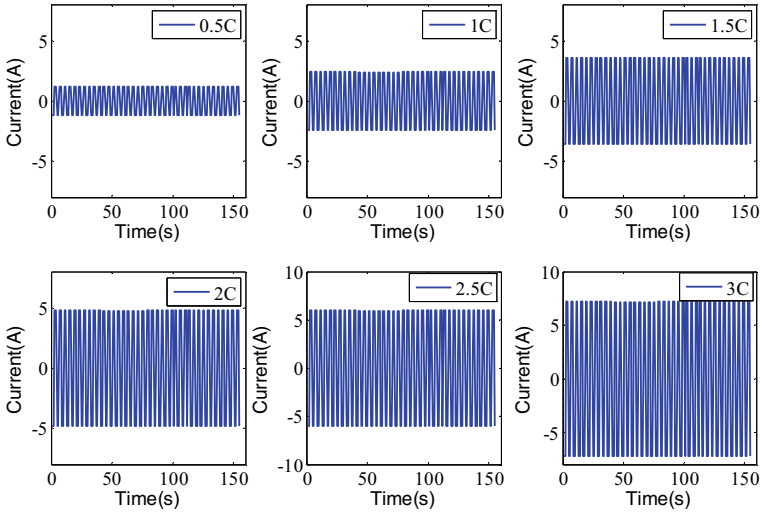


Fig. 3 The current versus time for the cell under the HPPC test

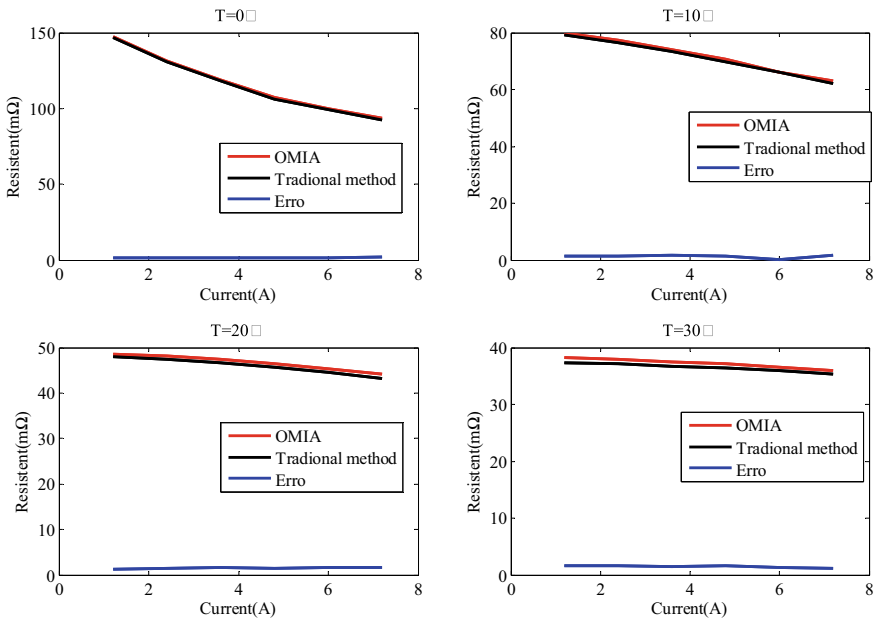


Fig. 4 Results of two methods at different temperature

5 Conclusion

In this study, a novel online model-based resistance algorithm (OMRA) has been developed to calculate the resistance of batteries at the different temperatures and amplitudes. The reliability and validity of the proposed method are verified by the mathematical model and experiments. The proposed strategy has the following obviously advantages:

(i) The proposed method is compatible with BV equation. (ii) The improved method is suitable for different temperatures and currents. It will not yield a large number of mini waves on the second scale as the traditional method. (iii) The proposed method requires no device and performs without interrupting. (iv) The presented method can calculate the battery resistance continuously and accurately, it can maintain the maximum error within 2 mΩ.

Further study will be focus on how to preheat the battery pack using this strategy. It aims to accurately obtain the resistance online and verify the applicability and validity of this method.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 51705480).

References

1. Zhang T, Zou Y, Zhang X et al (2020) Predictive eco-driving application considering real-world traffic flow. *IEEE Access* 99:1–1. <https://doi.org/10.1109/ACCESS.2020.2991538>
2. Wu H, Pang G, Choy K et al (2018) An optimization model for electric vehicle battery charging at a battery swapping station. *IEEE Trans Veh Technol* 67(2):881–895
3. Christophersen JP, Morrison J, Morrison W et al (2012) Rapid impedance spectrum measurements for state-of-health assessment of energy storage devices. *SAE Int J Passenger Cars Electron Electr Syst* 5(1):246–256
4. Kozłowski JD (2001) A novel online measurement technique for AC impedance of batteries and other electrochemical systems. In: Conference on applications & advances. IEEE, pp 257–62
5. Zou Y, Hu X, Ma H et al (2015) Combined state of charge and state of health estimation over lithium-ion battery cell cycle lifespan for electric vehicles. *J Power Sour* 273:793–803
6. Huang W, Qahouq JA (2014) An Online Battery Impedance Measurement Method Using DC–DC Power Converter Control. *IEEE Trans Industr Electron* 61(11):5987–5995
7. Howey DA, Mitcheson PD, Yufit V et al (2014) Online measurement of battery impedance using motor controller excitation. *IEEE Trans Veh Technol* 63(6):2557–2566
8. Qahouq JAA, Member S (2016) Online battery impedance spectrum measurement method. In: Applied power electronics conference & exposition. IEEE, pp 3611–5
9. Rahimi-Eichi H, Baronti F, Chow MY (2013) online adaptive parameter identification and state-of-charge coestimation for lithium-polymer battery cells. *IEEE Trans Industr Electron* 61(4):2053–2061
10. Zheng Y, Ouyang M, Han X et al (2018) Investigating the error sources of the online state of charge estimation methods for lithium-ion batteries in electric vehicles. *J Power Sour* 377:161–188

11. Lu L, Han X, Li J, Hua J, Ouyang M (2013) A review on the key issues for lithium-ion battery management in electric vehicles. *J Power Sour* 226:272–288
12. García-Hernández MT, Castilla J, González-Fernández CF, et al (1997) Application of the network method to simulation of a square scheme with Butler-Volmer charge transfer. *J Electroanal Chem* 424(1):207–212
13. Zhu J, Sun Z, Wei X et al (2016) Studies on the medium-frequency impedance arc for Lithium-Ion batteries considering various alternating current amplitudes. *J Appl Electrochem* 46(2):157–167
14. Ge H, Huang J, Zhang J et al (2016) Temperature-adaptive alternating current preheating of lithium-ion batteries with lithium deposition prevention. *J Electrochem Soc* 163(2):A290–A299
15. Li Z, Xiong R, Mu H, He H, Wang C (2017) A novel parameter and state-of-charge determining method of lithium-ion battery for electric vehicles. *Appl Energy* 207:363–371

A Research of Eco-driving Strategies Based on CVIS Considering Lane Change Behaviors



Jingyu Zhang, Chengcheng Xu, and Pan Liu

Abstract It's investigated that different lane change maneuvers will greatly influence the fuel consumption and emissions. Under the environment of cooperative vehicle infrastructure system (CVIS), on the basis of information interaction between vehicles before lane changing, to provide an eco-driving strategy for drivers can obviously reduce the fuel consumption of vehicles. On the foundation of analysis and summary of research findings from home and abroad in two fields of eco-driving and lane changing behavior, the paper proposes a lane change model applicable in the environment of CVIS by MATLAB, and by utilizing the traffic microcosmic simulation software of SUMO, compares the ecological benefits between the new model and traditional ones. It's concluded that the fuel consumption can be prominently reduced when following the proposed driving strategy. According to the sample data extracted from SUMO simulation, the average fuel consumption of a vehicle per kilometer saved by eco-driving strategy model under CVIS can be up to 19.86%.

Keywords Lane change · CVIS · Eco-driving · Fuel consumption · Emissions

1 Introduction

Nowadays, as the automobile industry is booming all over the world, expenditure of energy and environmental pollution problems have brought public attention on a large scale. Until the end of year 2019, the retention of vehicles in China has already reached 340 millions, whose high consumption of fuel and huge emissions have brought about great burdens to energy reserves and nature. In fact, the fuel consumption and emissions of vehicles depend on numerous factors, such as vehicle's own

J. Zhang (✉)

Department of Transportation, Southeast University, Nanjing 210018, Jiangsu, China

C. Xu · P. Liu

Key Laboratory of Urbain Intelligent Transportation, Southeast University, Nanjing 210018, Jiangsu, China

structure, vehicle load, frequency of use, maintenance quality and so on. Significantly, the driving style and use pattern of drivers could actually influence fuel consumption and emissions. For example, vehicles would probably stop, decelerate, change lanes, turn around when approaching a signalized intersection. These driving behaviors could cause an increase to fuel consumption and vehicle emissions. A number of drivers once participated in a field operational test in America, where drivers were given correspondent driving speeds to pass through intersections in the most eco-friendly way. The investigation showed that 20% of fuel could be saved on account of the optimized driving behaviors [1]. Therefore, the proposal of eco-driving strategies, namely optimizing driving behaviors or habits in order to reduce abrupt comportments like deceleration or acceleration, has significant meanings in response to sustainable development requirements.

The concept of eco-driving first reached American researchers. More precisely, researchers aimed to change driving style or habits by training drivers so as to decrease fuel consumption and emissions during their drive. Beusen with his team gathered driving information (mileage, number of revolutions per minute, position of the accelerator pedal and instantaneous fuel consumption) from 10 drivers during a period of 10 months and conducted a training process of 4 months. Their results presented a reduction of 5.8% [2] in average fuel consumption after training. However, it's worth noting that this investigation involved a small sample, not to mention that some drivers tended to fall back into their original drive habits. Similarly, the researches of Wahlberg [3] and Zarkadoula [4] both indicated that short-term impact of drive training could not possibly represent the real long-term impact.

Another research field of eco-driving leads to developing algorithms. For example, Mandava developed arterial velocity planning algorithms that give speed advice to drivers so that the probability of passing through a signalized intersection without stop could be maximized. In other words, the algorithms aimed to minimize the acceleration or deceleration rate. By using a stochastic simulation, the research showed a 12–14% decrease of energy or emissions in vehicles with speed advice [5].

Some other scholars considered that it's viable to control fuel consumption and emissions by improving traffic volume through intersections. Lee with his team proposed a CVIC (Cooperative vehicle Intersection Control) algorithm which seeks a safe maneuver for vehicles approaching intersections. It showed that under varying congestion conditions, CVIC algorithms significantly improved intersection performance. Meanwhile, it was claimed that emissions of CO₂ and fuel consumption were both reduced by 44% and 44% respectively [6].

Generally speaking, despite a number of researches proposed on eco-driving strategies until now, it remains still rare to find studies focused on eco-driving strategies considering lane change behaviors. Lane change behavior, beyond doubt, is regarded as a common driving comportment whether on urban roads or freeways. Therefore, to propose appropriate strategies serving vehicles with lane changing behaviors in an eco-friendly manner, has great significance for decreasing consumption, emissions and improving economic benefits. Taking advantage of the techniques of cooperative vehicle infrastructure system (CVIS), a vehicle is able to communicate

and cooperate with other vehicles before changing lanes. In this case, fuel consumption and emissions can be possibly reduced on account of optimized eco-driving strategies.

2 Modeling of Eco-driving Strategies for Lane Change Under CVIS

2.1 Differences of Lane Change Models Under CVIS

The development of CVIS techniques has promoted the realization of vehicle-to-vehicle and vehicle-to-infrastructure interaction and real-time information sharing. In general, under the conditions of CVIS, lane change models differ from the traditional ones in the following aspects:

1. Under the conditions of CVIS, vehicles could be informed of next actions (acceleration, deceleration, stop, turning around, etc.) of the forward vehicles based on real-time information sharing. In this case, the vehicle's car-following patterns would differ from the ones of traditional vehicles. More precisely, the safe following distance between vehicles or average headway would be reduced.
2. Under the conditions of CVIS, when target vehicle intends to change lanes, it could communicate with the vehicles forward and behind in order to negotiate and cooperate then come up with an optimal driving strategy.
3. Normally, as for lane change decision-making models, it's necessary to consider drivers' characteristics, including their driving style and habits, preferred lanes, aggressiveness, etc. Under CVIS, drivers usually would accept the advised driving strategies by on-board systems. Assuming that drivers follow the proposed strategies, the variety of personal characteristics would be greatly weakened.

2.2 Definition of Eco-friendly Lane Change Model

When building eco-driving models for lane changing vehicles, the essential of the efficiency of future modeling relies on the definition: what kind of lane change model is eco-friendly, which means by the way decrease of emissions and saving of energy? In fact, a number of vehicle behaviors could increase fuel consumption and emissions, namely idling, inappropriate speed, excessive load, bad driving habits (abrupt braking, sudden acceleration/deceleration), etc. [7]. In normal conditions, when a target vehicle intends to conduct lane change, it has to insert into a certain position on the target lane. When there are former or latter vehicles on the target lane, and sometimes the space is not visually large enough, the target vehicle and its latter vehicle on the target lane would probably decelerate even brake suddenly. This

type of comportment could not only increase fuel consumption and air pollution, but bring about a potential risk to human security.

In consequence, modeling of eco-driving strategies aims to optimize driving behaviors in an eco-friendly manner, in other words, to minimize impacts on vehicles involved during and after lane changing behaviors in order to avoid high consumption actions. In this case, the techniques of CVIS provide a foundation for the communication and cooperation between vehicles involved so as that vehicles could realize the lane changing behaviors steadily.

2.3 Modeling of Eco Lane Changing

Generally, when target vehicle intends to change to the target lane, the vehicles forward and behind on the target lane can be regarded as affected vehicles, shown in Fig. 1.

Assuming that target vehicle S, former vehicle F and latter vehicle L on the target lane travel at the speed of v_s , v_l and v_f , vehicle S intends to insert into a certain position between vehicle F and vehicle L. Before the action of lane changing happens, the original distance between target vehicle S and vehicle F is g_{of} , and between target vehicle S and vehicle L is g_{ol} . As long as the action of lane changing happens, the distance between vehicles involved will begin to change. In most cases, when the headway is visually large, vehicles involved travel at constant speeds during the lane changing behavior, the distance between target vehicle S and former vehicle F (g_f), latter vehicle L (g_l) during this period is noted as:

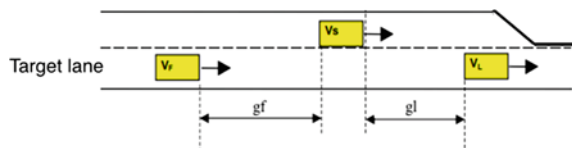
$$g_l = g_{ol} - v_s D_t + v_l D_t \tag{1}$$

$$g_f = g_{of} - v_f D_t + v_s D_t \tag{2}$$

If vehicles involved have to decelerate or accelerate during the lane changing behavior in case of small headways, the distance between target vehicle S and former vehicle F, latter vehicle L during the lane changing behavior is noted as:

$$g_l = g_{ol} - \left(v_s D_t + \frac{v_s D_t^2}{2} \right) + \left(v_l D_t + \frac{v_l D_t^2}{2} \right) \tag{3}$$

Fig. 1 Modeling diagram



$$g_f = g_{of} - \left(v_f D_t + \frac{v_f D_t^2}{2} \right) + \left(v_s D_t + \frac{v_s D_t^2}{2} \right) \quad (4)$$

Generally, if vehicles involved travel at constant speeds, the distance between target vehicle S and former vehicle F, latter vehicle L can both satisfy the safe spacing requirements of car-following models, shown as the expression below, this lane changing could be considered as a safe behavior without any influence on the other vehicles.

$$g_l \geq g_{min}, g_f \geq g_{min} \quad (5)$$

In fact, the minimum headway under the conditions of CVIS is smaller than the one in traditional car-following models. In consideration of security, in this research the minimum headway will take the tolerance value under the normal conditions. Sherif and Adel obtained the value of minimum headway in car-following patterns based on a great deal of data [8], which can be established as:

$$safe\ distance = l + V\delta + \frac{V^2}{2d_2} - \frac{V^2}{2d_1} \quad (6)$$

where

l vehicle length and minimum distance between consecutive vehicles (m);

V vehicle speed (m/s);

d_1 acceleration of former vehicle F (m/s^2);

d_2 acceleration of latter vehicle L (m/s^2);

δ constant of time (s).

Usually, the speeds and accelerations of former and latter vehicles in consecutive traffic flow are quite close. Additionally, time constant δ is always valued as 1 s, so the expression can be simplified as:

$$g_{min} = l + V \quad (7)$$

Using the calibration results from Sherif, this formula is shown as:

$$g_{min} = 8.6 + \bar{V} \quad (8)$$

where \bar{V} is average speed of vehicles on the segment (m/s).

2.4 Vehicle Fuel Consumption Model

Until now, scholars have put forward various types of fuel consumption models. For example, Barth brought about CMEM model based on fuel rate in 2005 [9], which considers numerous factors such as vehicle load, speed, road grade etc. Then

researcher Franceschetti [10] realized the calibration of model CMEM of Barth so as to guarantee the effective and practical use of the model.

The formula of CMEM model below represents the fuel rate of vehicle P_t (KW/s):

$$P_t = 0.5C_d\rho AV_t^3 + MV_t(g \sin \phi + gC_r \cos \phi + a_t) \quad (9)$$

where,

V_t vehicle speed (m/s);

a_t vehicle acceleration (m/s^2);

M vehicle mass (kg), which includes vehicle weight and load;

C_d air drag coefficient;

ρ air density (kg/m^3);

A vehicle frontal surface (m^2);

g gravity (kg/m^2);

ϕ road grade;

C_r friction coefficient.

The formula can be transferred to F_t (g/s):

$$F_t = \frac{\zeta}{\kappa} \left(kN_e V + \frac{P_t}{1000\epsilon\varpi} \right) \quad (10)$$

where

ζ air fuel ratio;

κ fuel calorific value (kJ/g);

k engine friction factor (kJ/rev/l);

N_e engine speed (rev/s);

ϖ vehicle drive system efficiency.

The formula can also be transferred to FR_t (l/s) to simplify the practical use in real life, in which ψ represents transformation factor:

$$FR_t = \frac{\zeta}{\kappa\psi} \left(kN_e V + \frac{P_t}{1000\epsilon\varpi} \right) \quad (11)$$

All the values of parameters above are shown as Table 1.

When considering the fuel consumption during a period of time T(s), it can be expressed as:

$$FC = \int_0^T FR_t \quad (12)$$

Therefore, based on the information above, a model of eco-driving strategies considering lane changing behaviors can be well built. The main idea is to validate if target vehicle S can achieve lane changing behavior with a constant speed safely and steadily, which means the minimum headway distance with former and latter vehicles

Table 1 Parameter values of CMEM model

Parameter	Physical meaning	Value
C_d	Air drag coefficient	0.7
ρ	Air density (kg/m ³)	1.204
A	Vehicle frontal surface (m ²)	3.912
g	Gravity (kg/m ²)	9.81
ϕ	Road grade	0
C_r	Rolling friction factor	0.01
ζ	Air fuel ratio	1
κ	Fuel calorific factor (kJ/g)	44
k	Engine friction factor (kJ/rev/l)	0.2
N_e	Engine speed (rev/s)	33
V	Engine displacement (l)	5
ε	Engine efficiency	0.4
ϖ	Vehicle drive system efficiency	0.9
ψ	Transformation factor (g/l)	737

on the target lane can be satisfied. Taking the minimum total fuel consumption in a certain distance (such as 1 km) as an optimization objective, an optimal speed strategy can be provided for vehicles involved in any conditions of original headways. Hence, the minimization of total fuel consumption for vehicles involved in lane changing behaviors can be achieved in this way.

3 Validation of Eco-driving Strategy Model

3.1 SUMO Vehicle Lane Changing Simulation

A simulation of the proposed eco-driving strategy model was conducted by the software SUMO. On the created one-way two lanes, two kinds of flow were input with different defined information such as maximum speed, maximum acceleration or deceleration, vehicle type, minimum headway distance, etc. By adjusting the speed distribution of the flow, vehicles can be induced to conduct lane changes in order to gain speed benefits. In this case, speed curves of vehicles involved (target vehicle S, former vehicle F and latter vehicle L) in the traveling distance of 1 km can be obtained after 20 groups of simulation. Figure 2 represents the speed curves of lane changing behaviors of 4 groups.

According to the simulation results above, it's evident that all the actions of lane changing behavior are accomplished within the first 10 s, when the speeds of target vehicle and latter vehicle L and former vehicle S are not stable. It indicates that the target vehicle often needs to change the speed with a large acceleration in order to

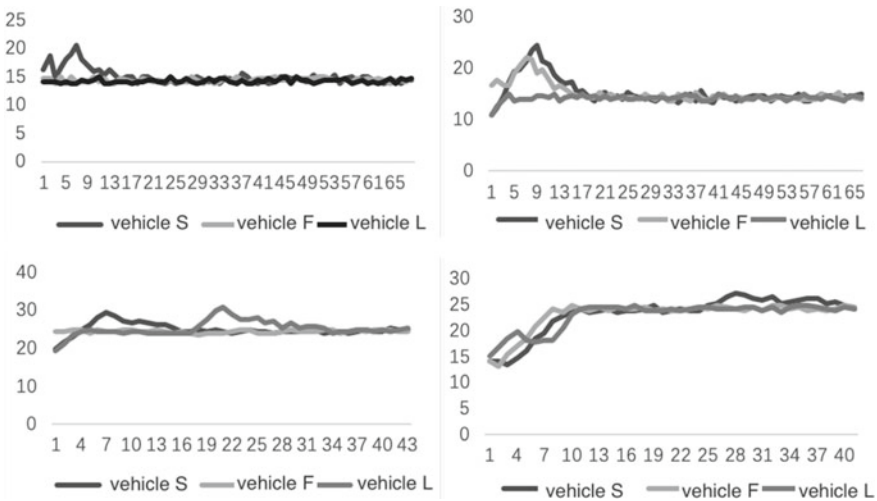


Fig. 2 Speed curves of lane changing behaviors simulated

perform the lane changing safely and smoothly. At the same time, the vehicles L and F on the target lane are also interfered, or even forced to change their own speeds to maintain a safe distance from the target vehicle S. In comparison with vehicles' speeds in the previous seconds of lane changing behavior, the speeds of vehicles involved tend to maintain relatively constant with an acceleration or deceleration close to zero.

As for the speeds of some vehicles varying after the lane changing behavior, there are many possible reasons concerned. The most likely one is that other vehicles in the lane have changed lanes at this time, which affects the speeds of vehicles.

3.2 Fuel Consumption Comparison

In order to verify the effectiveness of proposed eco-driving strategy model, it's necessary to compare the fuel consumption between the proposed model and traditional ones. Particularly, the speed curves from the simulation results are substituted as input to CMEM model so that the fuel consumption of each vehicle involved can be calculated.

In this research, the traditional model of lane changing follows the default of software SUMO. Additionally, in order to weaken the influence of traveling distances and the speed variations after conducting lane changing behaviors, the fuel consumption of models consider two different traveling distance: 0.5 km and 1 km respectively. The results of fuel consumption comparison are shown in Table 2.

From Table 2, it can be found that only in 3 groups the fuel consumption of the model of eco-driving strategies is greater than that of SUMO, while the other

Table 2 Fuel consumption comparison

No.	1 km			0.5 km		
	SUMO (l)	ECO (l)	%	SUMO (l)	ECO (l)	%
1	0.3349	0.3940	–	0.1973	0.1970	0.17
2	0.3527	0.2474	29.85	0.1829	0.1188	35.09
3	0.464	0.3501	24.55	0.2779	0.1749	37.05
4	0.3644	0.3562	2.24	0.2119	0.143	32.53
5	0.4166	0.3976	4.57	0.1973	0.1988	–
6	0.3931	0.3666	6.75	0.2114	0.1833	13.30
7	0.3801	0.2484	34.64	0.2025	0.1211	40.21
8	0.3577	0.3292	7.95	0.1687	0.1385	17.86
9	0.3875	0.3716	4.12	0.1971	0.1858	5.76
10	0.3758	0.2829	24.70	0.1983	0.1372	30.80
11	0.5085	0.3382	33.49	0.3265	0.1691	48.20
12	0.4545	0.334	26.51	0.2746	0.2991	–
13	0.4237	0.2938	30.66	0.2202	0.1316	40.22
14	0.4413	0.4362	1.16	0.2411	0.2181	9.55
15	0.4315	0.2902	32.74	0.2143	0.1288	39.89
16	0.4406	0.4096	7.04	0.2362	0.2048	13.30
17	0.551	0.2933	46.78	0.2368	0.2195	7.27
18	0.4343	0.2934	32.45	0.2228	0.2052	7.91
19	0.4507	0.4345	3.58	0.2228	0.2172	2.50
20	0.5051	0.286	43.39	0.3005	0.1257	58.19
Average	0.4234	0.3377	19.86	0.2271	0.1759	20.70

*Vehicle mass is assumed as 1500 kg

37 groups clearly present the effectiveness of the proposed model. Specifically, on account of eco-driving strategies, the average fuel consumption can be reduced by 19.86% within 1 km and 20.7% within 0.5 km. It can be noted that the model of eco-driving strategies has significant effects on saving fuel consumption in the process of lane changing.

3.3 Vehicle Emissions Comparison

Defra guidebook published in 2010 reveals the influence factors of vehicle emissions [11]. Taking CO₂ emissions as an example, the corresponding CO₂ conversion factor for vehicle fuel consumption is 3.808 g/km. Thus, the amount of CO₂ emitted by the simulation of SUMO lane changing model can be compared with the model of eco-driving strategies. Table 3 shows the CO₂ emissions of 20 groups of vehicles

Table 3 CO₂ emissions comparison

No.	1 km	0.5 km				
	SUMO (l)	ECO (l)	%	SUMO (l)	ECO (l)	%
1	0.3349	0.394	–	0.1973	0.197	0.17
2	0.3527	0.2474	29.85	0.1829	0.1188	35.09
3	0.464	0.3501	24.55	0.2779	0.1749	37.05
4	0.3644	0.3562	2.24	0.2119	0.143	32.53
5	0.4166	0.3976	4.57	0.1973	0.1988	–
6	0.3931	0.3666	6.75	0.2114	0.1833	13.30
7	0.3801	0.2484	34.64	0.2025	0.1211	40.21
8	0.3577	0.3292	7.95	0.1687	0.1385	17.86
9	0.3875	0.3716	4.12	0.1971	0.1858	5.76
10	0.3758	0.2829	24.70	0.1983	0.1372	30.80
11	0.5085	0.3382	33.49	0.3265	0.1691	48.20
12	0.4545	0.334	26.51	0.2746	0.2991	–
13	0.4237	0.2938	30.66	0.2202	0.1316	40.22
14	0.4413	0.4362	1.16	0.2411	0.2181	9.55
15	0.4315	0.2902	32.74	0.2143	0.1288	39.89
16	0.4406	0.4096	7.04	0.2362	0.2048	13.30
17	0.551	0.2933	46.78	0.2368	0.2195	7.27
18	0.4343	0.2934	32.45	0.2228	0.2052	7.91
19	0.4507	0.4345	3.58	0.2228	0.2172	2.50
20	0.5051	0.286	43.39	0.3005	0.1257	58.19
average	0.4234	0.3377	19.86	0.2271	0.1759	20.70

within 1 and 0.5 km. It can be seen that the eco-driving strategies based on CVIS can effectively reduce CO₂ and has a certain effect of energy saving and emission reduction.

4 Conclusion

This article elaborates on the differences of the proposed model and points out the positioning of the eco-driving strategies considering lane changing behaviors based on CVIS. The model of eco-driving strategies is built and the speed curve and driving information of the lane changing vehicle and other vehicles in the system are obtained through the simulation process. Then in comparison with the fuel consumption and CO₂ emissions of traditional lane changing model, the proposed model is well verified to have significant effects on fuel saving and emission reduction.

However, the optimization objective of this research only considers the indicator of vehicle fuel consumption. In the follow-up research, vehicle emissions such as CO, NO_x, HC and other gas emissions will be comprehensively considered to optimize the model of eco-driving strategies. In addition, when verifying the model of eco-driving strategies, the proposed model can be also applied into an appropriate simulation software, and the validity and the feasibility of the proposed model can be further verified through the simulation process and visualized operation.

References

1. Xia H, Boriboonsomsin K, Schweizer F et al (2012) Field operational testing of ECO-approach technology at a fixed-time signalized intersection. In: International IEEE conference on intelligent transportation systems. IEEE, pp 188–193
2. Beusen B, Broekx S, Denys T et al (2009) Using on-board logging devices to study the longer-term impact of an eco-driving course. *Transp Res Part D Transp Environ* 14(7):514–520
3. Wählberg AEA (2007) Long-term effects of training in economical driving: fuel consumption, accidents, driver acceleration behavior and technical feedback. *Int J Ind Ergon* 37(4):333–343
4. Zarkadoula M, Zoidis G, Tritopoulou E (2007) Training urban bus drivers to promote smart driving: A note on a Greek eco-driving pilot program. *Transp Res Part D Transp Environ* 12(6):449–451
5. Mandava S, Boriboonsomsin K, Barth M (2009) Arterial velocity planning based on traffic signal information under light traffic conditions. In: International IEEE Conference on Intelligent Transportation Systems. IEEE, pp 1–6
6. Lee J, Park B (2012) Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment. *IEEE Trans Intell Transp Syst* 13(1):81–90
7. Juntang C, Jianqiu Z (2011) Effective methods of reducing vehicle fuel consumption. *Chin Petrol Chem Stand Qual* 31(2):144–144
8. Abuelenin SM, Abul-Magd AY (2015) Effect of minimum headway distance on connectivity of VANETs. *AEU-Int J Electron C* 69(5):867–871
9. Barth M (2010) The comprehensive modal emission model (cmem) for predicting light-duty vehicle emissions. ASCE, pp 126–137
10. Franceschetti A, Honhon D, Woensel TV et al (2013) The time-dependent pollution-routing problem. *Transp Res Part B Methodol* 56(56):265–293
11. Defra (2010) Guidelines to Defra/DECC's GHG conversion factors for company reporting. http://www.sthc.co.uk/documents/DEFRA-guidelines-ghg-conversionfactors_2010.pdf

Research on Charging Station Planning Technology Based on Pure Electric Taxi SOC Driving Range in Northern Cold Area City



Haoran Liu, Bing Wang, Xuejing He, Junyou Sun, and Kuo Zhao

Abstract In this paper, through the analysis of the remaining mileage (SOC) in the operating characteristics of pure electric taxis in northern cold cities, the service range of charging stations is quantitatively analyzed and studied, so as to construct a geometric model for charging station planning, determine the constraints, and This paper proposes the shortest time service range model of charging stations based on SOC mileage, and solves the model using greedy algorithm to provide charging station planning technology for northern cold cities, so as to provide correlation for the promotion and application of pure electric taxis in northern cold regions Technical Support. Finally, based on the data related to the operation of pure electric taxis in Urumqi, the planning of the charging station in Urumqi was planned.

Keywords Northern cold city · Pure electric taxi · SOC · Driving range · Charging station · Planning technology

1 Background

The “Outline for the Construction of a Powerful Transport Country” clearly proposes to optimize the transportation energy structure, promote the application of new energy and clean energy, and promote the realization of electrification, new energy and cleanliness of all urban public transportation [1]. However, the current development layout of pure electric taxis in my country is mainly in the central and southern regions, and the development in the northern region is relatively slow, especially in some northern cold regions. The main reason for this uneven development phenomenon is the lack of technical theory and practical research on the promotion and application of pure electric taxis in the northern cold regions, and the technical research on the layout planning of pure electric taxi charging stations in cold cities is to promote pure electric taxis in the One of the key technologies for popularization and application in cold regions. At present, there are many studies on the layout of charging stations

H. Liu · B. Wang (✉) · X. He · J. Sun · K. Zhao
Xinjiang University, Urumqi, China

in conventional areas at home and abroad, but there is a lack of charging station planning research on the operating characteristics of pure electric taxis in northern cold areas. Therefore, the results of this research can be used for pure electric taxis and even pure electric vehicles in northern cold cities. The promotion and application of private cars provide certain scientific and technological theoretical support.

This article is based on the operation data of pure electric taxis in the cold city of Urumqi, with the vehicle operating characteristics, road network characteristics and temperature as the key factors affecting the service range of the charging station, to determine the constraints, based on the remaining power, the mileage and The service range of the charging station is constrained. The service range model of the charging station based on the remaining capacity SOC and the shortest mileage can be proposed. The model is solved by the greedy algorithm, and an optimal charging station planning scheme is finally obtained.

2 Analysis of Influencing Factors of Service Range of Urban Charging Stations in Northern Cold Regions

Based on the project “Analysis and Research on Operational Adaptability of Pure Electric Taxi in Urumqi City”, through analyzing the operating characteristics of pure electric taxis in Urumqi City, the main factors affecting the service scope of charging stations are pure electric vehicle model, temperature and road network.

(1) Vehicle type

At present, there are two types of pure electric taxis in Urumqi, A and B. The theoretical mileage of model A is 400 km, the theoretical mileage of model B is 300 km, the theoretical average mileage of model A is 5%, and the theoretical average mileage is 20 km. The driving range of 5% power is 15 km, as shown in Fig. 1. It can be analyzed from Fig. 1 that the greater the theoretical mileage of the vehicle’s 5% power, the greater the actual mileage of the vehicle’s 5% power.

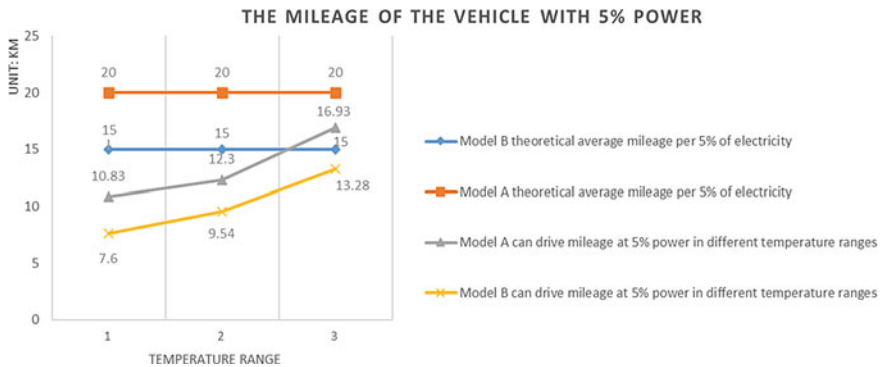


Fig. 1 The average mileage that the vehicle can drive for every 5% of electricity

(2) Temperature

As shown in Fig. 1, under the operating environment of the same model of pure electric taxis in different temperature ranges, the 5% of the vehicle's electricity can travel different distances.

Through a lot of data investigations on model A and model B, the actual driving of model A in each of the three different temperature ranges is 10.83 km, 12.3 km, and 16.93 km respectively. On average, the actual driving distance per 5% of electricity is 7.6, 9.54, and 13.28 km. At the same time, the actual mileage of the vehicle increases with the increase of the temperature range.

(3) Road network

Different road grades in the urban road network, the capacity of the road section, the driving speed of the road section, and the number of vehicles on the road section are different, and the travel time on each road section is different [2].

According to the data survey result of the 5% of the electric mileage of the pure electric taxi in the road network, the service scope of the charging station is analyzed to ensure that the charging station can meet the charging needs of vehicle types A and B in different operating environments, so With the minimum range of 5% of the 5% power below $-5\text{ }^{\circ}\text{C}$ of the vehicle type B as the effective service range of the charging station, the reachability of all pure electric vehicles to the charging station when remaining 5% of the power is guaranteed.

3 Charging Station Analysis

The charging station is the cornerstone of pure electric taxi operation and plays an irreplaceable role in the promotion and application of pure electric taxis. Research on the service scope of charging stations directly affects the number of charging stations. In order to plan and optimize the service range of the charging station, a geometric analysis method and a greedy algorithm are used to construct a service range model of the charging station with the shortest time based on the SOC mileage.

3.1 Basic Theory of Service Scope of Charging Station

The selection process of a pure electric taxi charging station is based on the SOC driving range. The service range of the charging station can be abstracted as a round range from the pure electric taxi to the charging station. Its size depends on the SOC driving range. Let D be the road node and S the charging station, then the theoretical geometry of the service range of the charging station is a circle with the charging station as the center. According to geometric knowledge, there are tangents, intersections or separations between different circle positions The three situations are shown in Fig. 2.

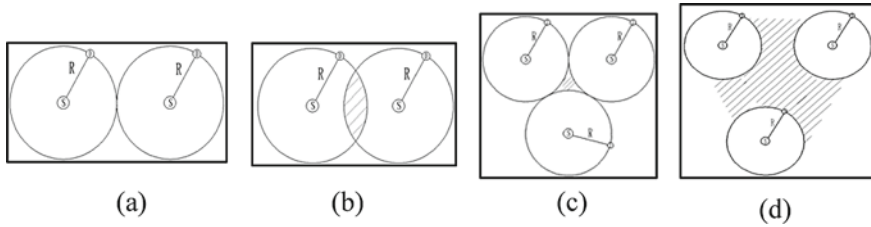


Fig. 2 Theoretical service scope model of the charging station

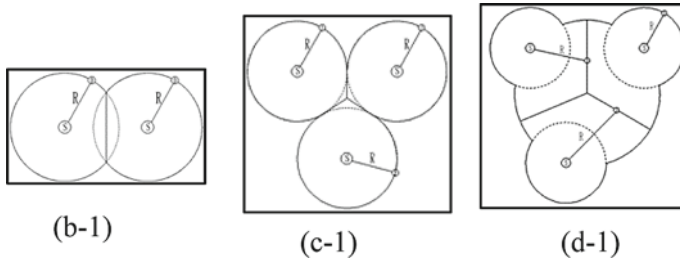


Fig. 3 Distribution form of the theoretical service range of the charging station after correction

Figure 2 shows the existence mode of the theoretical service range of the charging station (a) The figure shows that the service ranges of the two charging stations are tangent, which is an optimal service range mode of the charging station. (b) The figure shows that the service areas of the two charging stations form an intersecting area. In the intersecting area, pure electric taxi drivers can reach any one of the charging stations to charge, forming an overlap in the service scope of the charging stations, resulting in charging resources waste. (c) The figure shows that the service range of multiple charging stations is tangent, resulting in a small area not covered by the service range. (d) The figure shows that the service ranges of multiple charging stations are separated, indicating that the distribution of charging stations is insufficient to form a large area that is not covered by the service.

Based on the analysis of b–d in Fig. 2, the service range of the charging station has overlapping or uncovered coverage in the actual road network, so it is necessary to the service radius is revised and adjusted, and the shadow area is evenly divided to correct the service range of the charging station to ensure the rationality of the service range of the charging station, as shown in Fig. 3.

3.2 Model Building

By optimizing the geometric model of the distribution of charging stations, the service range of the charging station is optimized, with the charging station as the center, the

SOC mileage of the pure electric taxi is the service radius of the charging station, and the radius based on the SOC of the pure electric taxi is built. The shortest time model of the service range ensures that the SOC mileage of each vehicle in the road network can safely reach each charging station to receive charging service, and the greedy algorithm is introduced to the charging service range of the charging station based on the pure electric taxi SOC mileage. The shortest time model of the radius of service range is solved.

3.2.1 Objective Function

When driving a pure electric taxi on the road network, it should be ensured that the SOC can travel to the charging station for charging and avoid stranding. Therefore, the charging station should theoretically cover the road network to ensure the accessibility of the pure electric taxi. The determination of the service range of the charging station is very important for the operation of the pure electric taxi, so the service range model of the charging station based on the shortest time that the SOC can travel is constructed.

According to the regression resistance model proposed by domestic scholars [3]:

$$T = t_0 \left[1 + K_1 \left(\frac{Q_1}{C_1} \right) + K_2 \left(\frac{Q'_1}{C'_1} \right) \right] \tag{3.1}$$

where: T is the travel time of the road segment, h ; t_0 is the travel time of the free travel of the road segment, h ; K_1 and K_2 are regression parameters, which can be estimated by the maximum likelihood estimation method [4], both are greater than 0; Q_1 is the traffic volume of motor vehicles and Q'_1 is the traffic volume of non-motor vehicles in the road section, pcu/h; C_1 is the traffic capacity of motor vehicles and C'_1 is the traffic capacity of non-motor vehicles in the road section, pcu/h.

Build a service range based on the mileage of pure electric taxi SOC, and determine the time for pure electric taxi to reach the charging station as T_d :

$$T_k = t_{0k} \left[1 + K_1 \left(\frac{Q_k}{C_k} \right) + K_2 \left(\frac{Q'_k}{C'_k} \right) \right] \tag{3.2}$$

In the formula: T_k is the travel time of the k -th section of the road, h ; t_{0k} is the free travel time of the k -th section of the road, h ; K_1 and K_2 are regression parameters, both are greater than 0; Q_k is the traffic volume of motor vehicles and Q'_k is the traffic volume of non-motor vehicles in section k , pcu/h; C_k is the traffic capacity of motor vehicles and C'_k is the traffic capacity of non-motor vehicles in section k , pcu/h.

$$T_d = \sum_{k=1}^{m_h} T_k = \sum_{k=1}^{m_h} \left\{ t_{0k} \left[1 + K_1 \left(\frac{Q_k}{C_k} \right) + K_2 \left(\frac{Q'_k}{C'_k} \right) \right] \right\} \tag{3.3}$$

In the formula: T_d is the time taken to charge the pure electric taxi, h; m_h is the m -th path for the pure electric taxi to go to the charging station for charging, and the path is composed of h road segments.

The vehicle density [5] of the road section is:

$$\rho = n/L \tag{3.4}$$

$$Q = v\rho \tag{3.5}$$

where: ρ is the vehicle density of the road section, pcu/km; n is the number of vehicles on the road section, pcu; L is the length of the road section, km; Q is the traffic volume of the road section, pcu/h; v is the design speed of the road section.

You may have:

$$Q_k = v_k\rho_k = v_k \frac{n_k}{L_k} \tag{3.6}$$

$$Q'_k = v'_k\rho'_k = v'_k \frac{n'_k}{L_k} \tag{3.7}$$

where: ρ_k is the motor vehicle density and ρ'_k is non-motorized vehicle density on the k -th section of the road, respectively, pcu/km; n_k are the number of motor vehicles and n'_k are non-motorized vehicles on the k -th section of the road respectively Motor vehicle, pcu; L_k is the length of the k th section of the road, km; v_k is the design speed of the motor vehicle and v'_k is the design speed of non-motorized vehicle on the k -th section of the road.

Travel time of pure electric taxi to charging station:

$$T_d = \sum_{k=1}^{m_h} \left\{ t_{0k} \left[1 + K_1 \left(\frac{v_k n_k}{C_k L_k} \right) + K_2 \left(\frac{v'_k n'_k}{C'_k L_k} \right) \right] \right\} \tag{3.8}$$

3.2.2 Constraints

According to the model assumption, it is necessary to ensure that the layout of the charging station covers the entire area, and also to ensure that the overlapping area between the service areas of adjacent charging stations is minimal. Due to the different road resistance between the urban core area and the urban core area, the service of the urban core area charging site is caused. The range is different from the service range of the charging station outside the city core area, so the road network is simplified as shown in Fig. 4. Under the geometric analysis theory, the distance between $S_1 \sim S_i$ is the largest, so that S_1 and S_i are adjacent to each other. The service scope of the charging station covers all roads in the road network tangentially.

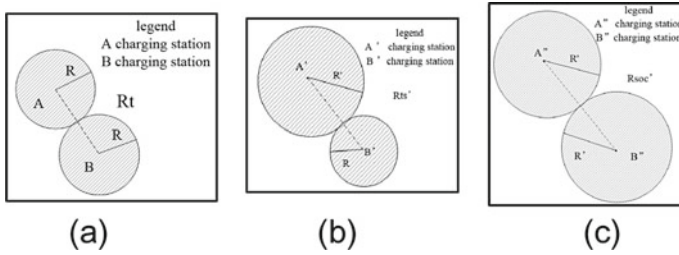


Fig. 4 Sketch map of the service area covering the smallest circle of all urban roads

Therefore, the constraint conditions for determining the objective function of Formula (3.3) are:

$$\min T_d = \sum_{k=1}^{m_h} \left\{ t_{0k} \left[1 + K_1 \left(\frac{v_k n_k}{C_k L_k} \right) + K_2 \left(\frac{v'_k n'_k}{C'_k L_k} \right) \right] \right\} \quad (3.9)$$

$$\text{s.t. } L_{m_h} = \sum_{k=1}^N L_k \leq L_{soc} \quad (3.10)$$

$$R_i = R \quad (3.11)$$

$$R'_i = R' \quad (3.12)$$

$$R_i + R_{i+1} \leq R_t \quad (3.13)$$

$$R_i + R'_{i+1} \leq R'_{ts} \quad (3.14)$$

$$R'_i + R'_{i+1} \leq R'_{soc} \quad (3.15)$$

$$N \in M$$

where R is the service radius of the charging station in the city's core area, km; R' is the service radius of the charging station determined by the mileage of the SOC under certain conditions outside the city's core area, km; R_t is between two adjacent charging stations in the city's core area Coverage service range, km; R'_{ts} is the service range covered between the city's core area and two adjacent charging stations outside the city core area, km; R'_{soc} is the service range between the two adjacent charging stations outside the city's core area, km. $G(N, M)$ is the topological network of charging stations, M is the road segment set, and N is the charging station set.

In summary, the expression of the objective function is shown in the Formula (3.9), which represents the shortest time for the pure electric taxi to reach the charging station; the expression of the constraint condition is shown in the Formula (3.10), the path for the pure electric taxi to charge the charging station The mileage should be less than the mileage that can be driven by 5% SOC; the constraints are expressed in Formulas (3.11) and (3.12), which represent the service radius of charging stations in the city’s core area and the service radius of charging stations outside the city’s core area; For the expression of constraints, see Formulas (3.13), (3.14) and (3.15), limit the service radius of each charging station, and minimize the intersection area of the service range of the charging station, restrict all pairs of charging stations All roads are covered by services.

3.3 Model Solution

In order to solve the above model, first determine the charging station set N , the road segment set M for the charging station service range determination of the shortest time based on the SOC mileage of each charging station in the road network. The model is solved by the greedy algorithm, and the overall optimal constraint conditions Iteratively calculate the respective optimal service service range for each charging station [6], and finally obtain the charging station service range of all charging stations based on the SOC-driving distance in the shortest time. The solution process is as follows:

Step 1: First label all charging stations in the road network topology map S_i ($i = 1, 2, 3 \dots$), and then label each road node D_j ($j = 1, 2, 3 \dots m$);

Step 2: Initially select the actual vehicle speed, actual vehicle number, and basic traffic capacity corresponding to the sections S_1 to D_j , the passing urban roads L_k , L_k and calculate the time from S_1 to each road node D_j according to Formula (3.10) T_{dk} ($k = 1, 2, 3 \dots$);

$$T_{dk} = \sum_{k=1}^{m_h} \left\{ t_{0k} \left[1 + K_1 \left(\frac{v_k n_k}{C_k L_k} \right) + K_2 \left(\frac{v'_k n'_k}{C'_k L_k} \right) \right] \right\} \tag{3.16}$$

Step 3: Set N_i and M are established, and the T_{dk} obtained in Step 2 is discriminated according to Formula (3.11), and the corresponding road node D_j less than the constraint condition L_{soc} is placed in the set N_i . Add 1 to continue calculation;

$$N_i = \{D_j\} \tag{3.17}$$

Step 4: Repeat Step 2 and Step 3 to determine whether D_j has calculated all nodes and S_i has calculated all charging stations. When D_j and S_i are both calculated, assign the obtained N_i to set M , and repeat steps 2 and 3 above when D_j and S_i are not calculated;

$$M = \{N_i\} \tag{3.18}$$

Step 5: Determine the circle center distance between each adjacent charging station in the set M in step four, when the adjacent circle center distance meets the Formulas (3.13), (3.14), (3.15) under different conditions. At this time, all the road nodes in the set M are connected with curves to obtain the charging station service range of the charging station based on the shortest time that the SOC can drive the mileage. When the adjacent circle center distances do not satisfy the Formulas (3.13), (3.14), (3.15), the service range of the charging station is corrected by the basic theory of the service range, and finally the charging station based on the SOC mileage can be obtained. The service range of the charging station in the shortest time.

Figure 5 shows the flow of charging station service range based on the greedy algorithm to solve the shortest time of charging station based on SOC mileage.

Thus, the charging station service range of the charging station based on the shortest time of the SOC mileage can be established. Actually, the service range of each charging station can be expressed as: S is the charging station, D is the road node, then the charging station's service range of the charging station based on the SOC mileage can be shown in Fig. 6.

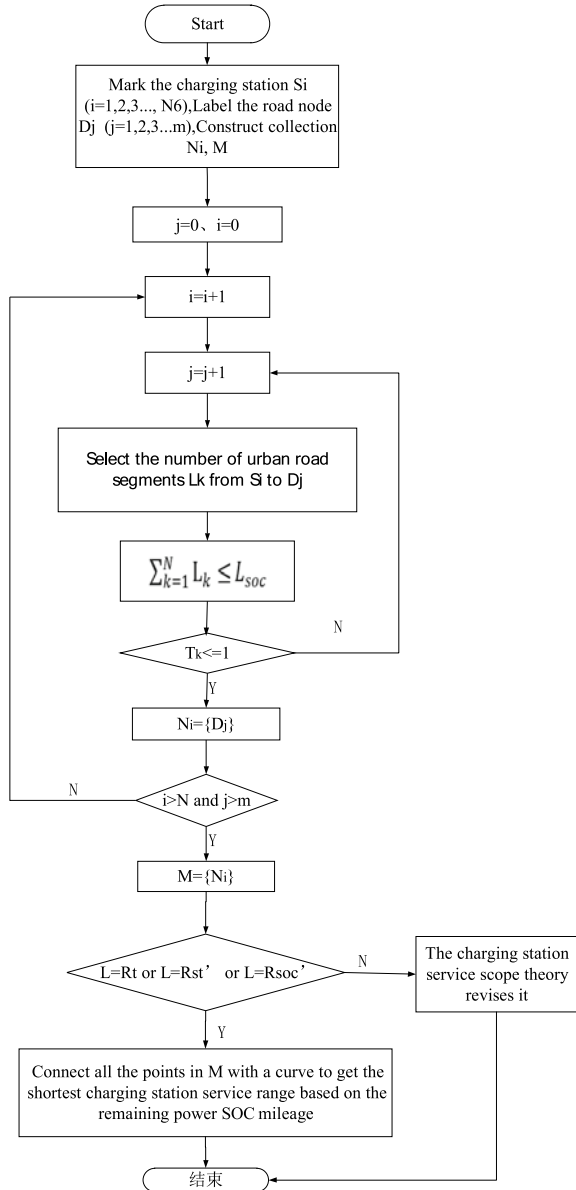
4 Application Analysis

Based on the project "Urumqi City Pure Electric Taxi Operation Suitability Analysis Research", analysis The data shows that the 5% SOC of pure electric taxis in Urumqi can drive 7.6 km, and the service radius of the charging stations in the core areas of major cities such as the provincial capital that actively promotes the area is less than 2 km []. Based on the solution process of the shortest charging station service range model based on the SOC driving mileage constructed above, the service range of the charging station in Urumqi is optimized. According to the design speed, basic traffic capacity, road length, etc. of various grades of roads in Urumqi [7], $R = 2$ km, $R' = 7.6$ km, solve the Urumqi charging station planning, as shown in Fig. 7.

5 Conclusion

By analyzing several existing forms of service scope of the charging station, the service scope is optimized. Construct the service range model of the charging station based on the shortest time of SOC driving mileage, obtain the constraint conditions by geometric analysis of the actual measurement of the satellite image map, and finally solve the model through the greedy algorithm, taking Urumqi as an example, the solution is severe cold. The planning of charging stations in the region provides certain planning techniques.

Fig. 5 Flowchart of solving the charging station service scope based on the shortest time the SOC can drive the charging station



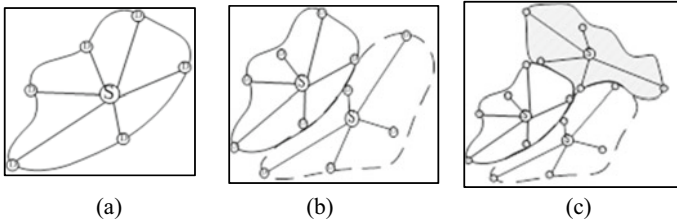
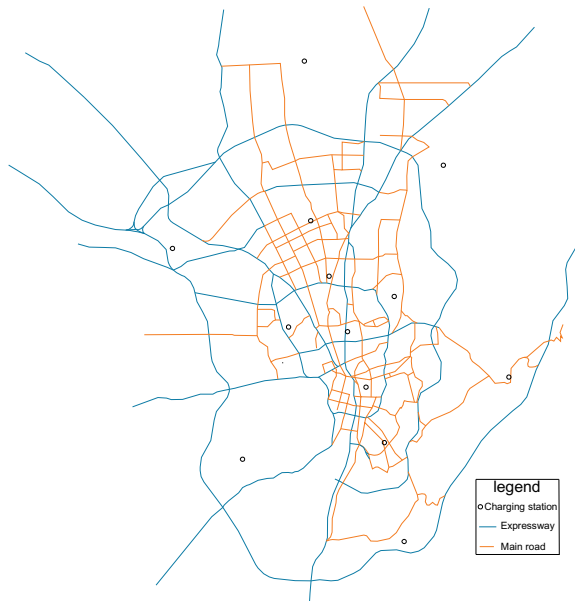


Fig. 6 Charging station service range of charging station based on SOC mileage

Fig. 7 Planning of charging stations in Urumqi



References

1. The Central Committee of the Communist Party of China issued the “Outline for Building a Powerful Country”, 2019.9
2. Wei W, Xiucheng G (2011) Traffic engineering. Southeast University Press, Nanjing
3. Wei W, Xuewu C (2008) Transportation planning. People’s Communications Press, Beijing
4. Wenzhou J, Jie Z, Xuli Z (1996) Maximum likelihood calibration method of road resistance function. Highway Transp Sci Technol 13(4):24–28
5. Xu J, Chen X (2008) General introduction to traffic engineering. People’s Communications Press, Beijing
6. Cao X (2018) Intelligent dormitory allocation method based on greedy algorithm. Comput Modern 01:23–26+31
7. Jianjun W, Baojie Y (2004) Traffic investigation and analysis. People’s Communications Press, Beijing

Model Construction and Simulation Analysis of Vehicle Collision on Steep Road Section in Mountainous Area



Xuezhao Fang, Bing Wang, Haoran Liu, and Kuo Zhao

Abstract Mountainous areas account for a large proportion of China's land area. Mountain highway is an important part of our country's highway system. The line shape of mountain highway is complex. The driver's line of sight is affected by the slope when driving in this section, and the safe sight distance is insufficient, which can easily cause vehicles to collide with traffic accidents time and rollover in high subgrade section at the same. This paper analyzes the influence of the steep slope on the driver's sight distance, and quantitatively analyzes the relationship between the sight distance and the slope during the driving process of the driver in the steep slope section, so as to provide a reference for ensuring the safe sight distance of the steep slope section of the mountain road. The dynamic analysis of vehicle collision on steep slope is carried out, and the relationship between vehicle speed and collision deformation is studied. The dynamic simulation of vehicle collision is carried out by Pc-Crash simulation software, which provides suggestions for the safe driving speed of vehicles on steep slope sections of mountain roads so as to ensure the driving safety of vehicles on steep slopes in mountainous areas. It can reduce the property loss caused by vehicle collision in mountainous steep slope section and better protect the driver's personal safety.

Keywords Mountain road · Steep road section · Driving sight distance · Kinetic analysis · Pc-Crash simulation

1 Introduction

China has a large mountain area, which accounts for about $2/3$ of the total land area [1]. The construction of mountain roads is greatly restricted by the geological terrain. Some sections are high subgrade sections with large road slope, which causes great inconvenience to the safe passage of vehicles. Meanwhile, the vehicles are vulnerable to the impact of rainfall, fog and other special weather conditions in the mountain

X. Fang · B. Wang (✉) · H. Liu · K. Zhao
Xinjiang University, Urumqi 830047, China

area during the driving process, which is very easy to cause traffic accidents such as vehicle collision and rollover [2]. Through the summary and analysis of the overall traffic accident statistics results, from the perspective of accident form, the number of rollover or collision with roadside obstacles of vehicles in high subgrade section accounts for 50% of the total number of accidents, from the perspective of accident area, the proportion of accidents in mountainous and hilly areas accounts for 44.8%, among which the three provinces of Yunnan, Guizhou and Tibet are more prominent [3].

Due to the excessive number of vehicle collision accidents on the steep slope of mountain highway, which results in large property loss and more casualties, the causes of vehicle collision accidents on the steep slope of mountain highway are analyzed in-depth at home and abroad. Feng Yutao et al. Analyzed the causes of the common diseases in the mountain steep slope sections through the research on the mountain steep slope sections, put forward the relevant technologies to ensure the safety of the mountain steep slope sections, and provided the ideas to solve the problems for the subsequent safety research of the steep slope sections [4]. Based on the reliability theory, Xu Tao et al. constructed function functions for vehicle side slip and insufficient sight distance and quantitatively analyzed the effect of vehicle speed control on reducing the probability of side slip [5]. Dissanayake et al. used logit model to analyze the influencing factors of the severity of accidents on both sides of the single vehicle rush out road section from the perspective of people, vehicles, roads and environment, and found a number of factors significantly related to death accidents [6]. Based on the historical roadside accident data, Bambach et al. compared and analyzed the predicted severity of the collision between motor vehicles and various roadside barriers, and found that roadside barriers have significant effect in reducing accident damage [7]. Roque et al. found that the three factors, slope gradient, roadside obstacles and rolling of vehicles have a very serious impact on the severity of the accident after they rush out of the road [8]. Experts at home and abroad have carried out in-depth research on the influencing factors of vehicle collision on the steep slopes of mountain roads, and analyzed the influence degree of different influencing factors on vehicle collision. However, there are few studies on the relationship between slope and driving sight distance, lack of quantitative analysis on the severity of collision after vehicle collision. The follow-up researchers lack of intuitive feelings on the results of vehicle collision.

In this paper, based on the analysis of the risk factors of driving on the steep slope of mountain highway, the relationship between driving sight distance and steep slope is studied, and the safe driving sight distance of drivers on the steep slope is further studied. According to the collision dynamic model, the dynamic analysis of vehicle collision accidents on the steep slope is carried out, and the relationship between vehicle collision deformation and vehicle speed is constructed. At the same time, through the Pc-Crash traffic accident simulation software, a three-dimensional collision model is built to quantitatively analyze the vehicle collision, which provides the following researchers with an intuitive experience of vehicle collision. This paper is supported by the natural science foundation of Xinjiang University autonomous region. The project number is 2018d01c054.

2 Analysis of Sight Distance of Drivers on Steep Slope Section of Mountain Highway

When driving on mountain roads, the main factors that affect vehicle collision are vehicle power and sight distance, among which the main factor is the driver’s sight distance. The construction of the slope of the mountain road ramp needs to be carried out in accordance with the road engineering specifications. The relationship between the driving speed and the slope length and slope is shown in Tables 1 and 2 [9].

When the highway with design speed of 120, 100 and 80 km/h is limited by terrain conditions or other special circumstances, the maximum longitudinal slope can be increased by 1% through technical and economic demonstration; in highway reconstruction, the section with design speed of 40, 30 and 20 km/h using the original highway can be increased by 1% through technical demonstration.

In high altitude area, the driving force and climbing ability of vehicles are reduced due to the influence of external environment. The maximum longitudinal slope value of highways at all levels above 3000 m above sea level shall be reduced in accordance with the provisions of Table 3. If the maximum longitudinal slope is less than 4%, 4% shall still be used (Table 4).

When the driver is driving in the steep mountain section, his sight is blocked by the hillside slope, so it is difficult to observe the road condition information of the road ahead to ensure the safe sight distance. The analysis of driver’s sight distance on steep road section is shown in Fig. 1.

It can be seen from Fig. 1 that when the driver drives the vehicle on both sides of the slope top of the steep road section, the line of sight is blocked by the hillside body, so it is difficult to observe the information of the road ahead. It is difficult to take measures in time when the vehicle finds the opposite vehicle on the slope top, which is easy to cause the vehicle collision accident. The vehicle is easy to cause the vehicle rollover accident due to sideslip when the steep road section is a high

Table 1 Minimum slope length of highway (m)

Design speed (km/h)	120	100	80	60	40	30	20
Minimum slope length	300	250	200	150	120	100	60

Table 2 Maximum longitudinal slope of highways at all levels

Design speed (km/h)	120	100	80	60	40	30	20
Maximum longitudinal slope (%)	3	4	5	6	7	8	9

Table 3 Impairment of plateau longitudinal slope break (%)

Altitude (m)	3000–4000	4000–5000	5000 above
Write down	1	2	3

Table 4 Minimum radius and length of vertical curve of highway (m)

Design speed (km/h)		120	100	80	60	40	30	20
Radius of convex vertical curve	General value	17,000	10,000	4500	2000	700	400	200
	Limit value	11,000	6500	3000	1400	450	250	100
Minimum length of vertical curve		100	85	70	50	35	25	20

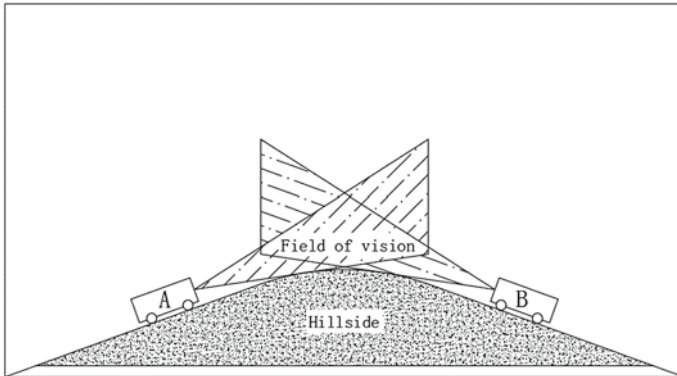


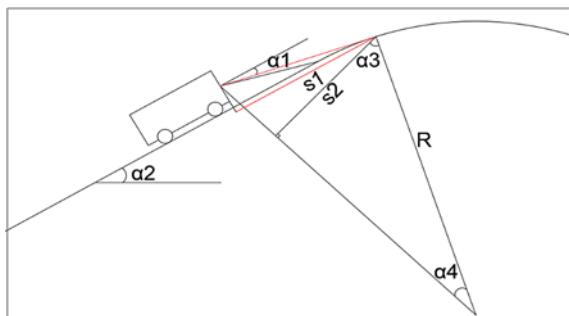
Fig. 1 Schematic diagram of driving sight distance of drivers on steep slopes

subgrade section. The upper and lower visual field width of the driver during driving is 15° [10].

Vehicle sight distance refers to the longest distance that a driver can continuously see a certain height of obstacles, traffic facilities and road markings in front of the road from the normal driving position in the normal driving process. The distance here refers to the length measured along the lane centerline. The visual field height of car is 1.08 m, and that of truck is 2.4 m [11].

It can be seen from Fig. 2 that the angle between the tangent line of the driver's line of sight and the vertical curve of the ramp and the line of sight is α_1 , which is also the boundary of the driver's down view. The gradient of the steep slope section

Fig. 2 Diagram of sight distance and slope analysis



is α_2 . The angle between the driver's line of sight and the tangent point of the ramp to the center point of the curve between the driver and the ramp and the tangent point to the center point of the curve is α_3 . The included angle between the line and the center point is α_4 . The driver's sight distance is S_1 . The distance between the driver's sight line and the tangent point of the ramp and the line between the driver's position and the center point is S_2 .

From the analysis, α_1 can be obtained from the actual measurement data, α_2 is the gradient of steep road section, and:

$$\alpha_1 + \alpha_3 \approx 90^\circ \tag{1}$$

$$\alpha_1 = \alpha_4 \tag{2}$$

$$\sin \alpha_4 = \frac{s_2}{R} \tag{3}$$

In the actual driving process, due to the small angle of vision, the driver can get:

$$s_1 \approx s_2 \tag{4}$$

From this, we can get:

$$s_2 = R * \sin \alpha_4 \tag{5}$$

It refers to the sight distance when the vehicle is on the steep slope.

From the above analysis, it can be concluded that with the decrease of the distance between the vehicle and the top of the slope and the driver's sight distance is gradually less. The driver's sight distance is the smallest when the driver's head up line is tangent to the top of the slope. As the vehicle continues to drive, the driver's sight distance is gradually increased. The driver's sight distance is the largest when the vehicle reaches the top or bottom of the slope. The information of the road ahead can be observed when the driver's sight distance is large. At this time, vehicle collision accidents are not easy to occur. It is difficult to observe the information of the road ahead when the driver's sight distance is small, and collision accidents are easy to occur. Through further analysis of vehicle collision accidents on steep slopes, the relationship between vehicle collision deformation and vehicle speed can be obtained, which has a good promoting effect on the study of how to reduce collision loss and deal with it quickly after collision accidents.

3 Dynamic Analysis and Model Construction of Vehicle Collision on Steep Slope

3.1 Vehicle Collision Dynamics Analysis on Steep Road Section

The driving speed and the combined external force are the main factors affecting the deformation when the vehicle is driving on steep slope. The dynamic analysis of the vehicle in collision accident is convenient to find the relationship between the vehicle deformation and its related influencing factors. For the vehicle in collision, it can more accurately determine the running state of the vehicle before the collision accident, which is conducive to the determination of the accident. At the same time, it can explore the vehicle collision loss from the source, in order to better formulate relevant protective measures, so as to reduce property loss and casualties. Orthogonal decomposition is carried out for the vehicle speed before collision accident in the lane, as shown in Fig. 3.

As shown in Fig. 3, take the vehicle center of gravity as the coordinate origin, take the horizontal direction as the x-axis direction, and take the vertical direction as the z-axis direction. Suppose that car A and car B are traveling in the opposite direction on the steep road section, and the initial speed of car a is v_{A0} . Its projection component in the x-axis direction is v_{A0X} . Its projection component in the z-axis direction is v_{A0Z} . Its angle with the x-axis direction is α_{A0} . The initial speed of vehicle B is v_{B0} , whose projection component in the x-axis direction is v_{B0X} . Its projection component in the z-axis direction is v_{B0Z} . The angle between b_0z and x-axis is β_{B0} . Available:

$$v_{A0} = \sqrt{(v_{A0X})^2 + (v_{A0Z})^2} \tag{6}$$

$$v_{A0X} = v_{A0} \cos \alpha_{A0} \tag{7}$$

Fig. 3 Analysis diagram of vehicle speed before collision in steep slope section

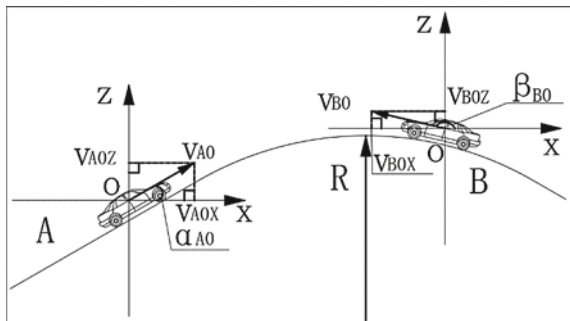
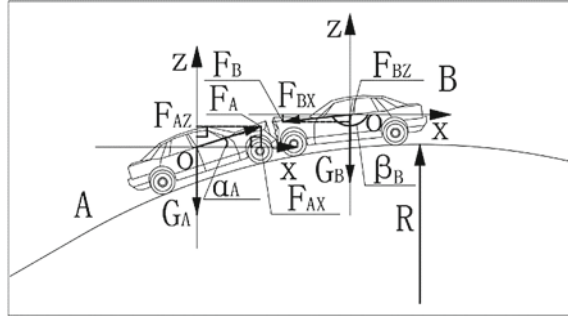


Fig. 4 Analysis diagram of vehicle speed before collision in steep slope section



$$v_{A0Z} = v_{A0} \sin \alpha_{A0} \tag{8}$$

$$v_{B0} = \sqrt{(v_{B0X})^2 + (v_{B0Z})^2} \tag{9}$$

$$v_{B0X} = v_{B0} \cos \beta_{B0} \tag{10}$$

$$v_{B0Z} = v_{B0} \sin \beta_{B0} \tag{11}$$

It is assumed that the initial vehicle speed of car A is v_{A1} . Its projection component in the x-axis direction is v_{A1X} . Its projection component in the z-axis direction is v_{A1Z} . Its angle with the x-axis is α_{A1} . The initial speed of vehicle car B is v_{B1} . Its projection component in the x-axis direction is v_{B1X} . Its projection component in the z-axis direction is v_{B1Z} . Its angle with the x-axis direction is β_{B1} . The corresponding relationship can be obtained from the above analysis.

Orthogonal decomposition is carried out for the force on the vehicle running in the lane before the collision accident, as shown in Fig. 4.

As shown in Fig. 4, take the vehicle center of gravity as the coordinate origin, take the horizontal direction as the x-axis direction, and take the vertical direction as the z-axis direction. Traction force of vehicle a is F_A . Its projection in the x-axis direction is F_{AX} . Its projection in Z direction is F_{AZ} . Its angle with the x-axis direction is α_A . Its own gravity is G_A . Traction force of vehicle B is F_B . Its projection in the x-axis direction is F_{BX} . Its projection in Z direction is F_{BZ} . Its angle with the x-axis direction is β_B . Its own gravity is G_B . From the above analysis, we can get:

$$F_A = \sqrt{(F_{AX})^2 + (F_{AZ})^2} \tag{12}$$

$$F_{AX} = F_A \cos \alpha_A \tag{13}$$

$$F_{AZ} = F_A \sin \alpha_A = G_A \tag{14}$$

$$F_B = \sqrt{(F_{BX})^2 + (F_{BZ})^2} \quad (15)$$

$$F_{BX} = F_B \cos \beta_B \quad (16)$$

$$F_{BZ} = F_B \sin \beta_B = G_B \quad (17)$$

3.2 Model of Vehicle Collision Deformation and Speed on Steep Slope

In the process of vehicle collision, part of the deformation of the vehicle occurs. After the collision, the deformation part of the vehicle is not completely recovered. Therefore, vehicle collision is not an elastic collision, nor a completely inelastic collision, but an inelastic collision. The elastic recovery coefficient [12] in collision is expressed by ε , which is defined as follows:

$$\varepsilon = \frac{v_{B1} - v_{A1}}{v_{A0} - v_{B0}} \quad (18)$$

where v_{A0} is the velocities of car A at the moment of contact before collision (m/s), v_{B0} is the velocities of car B at the moment of contact before collision (m/s). v_{A1} is the velocities of car A at the moment of contact after collision (m/s), v_{B1} is the velocities of car B at the moment of contact after collision (m/s). The recovery coefficient is approximately zero when the effective collision speed reaches more than 20 km/h. The structure of modern vehicles is basically the same and the difference of vehicle materials is not big and the plastic deformation difference between vehicles is not big [13].

The process of vehicle collision is analyzed from the perspective of kinetic energy conservation. It is assumed that the changes of kinetic energy before and after vehicle collision are all converted into the deformation energy of vehicle, namely:

$$\Delta E_{def} = \Delta E \quad (19)$$

where the ΔE_{def} is the deformation energy of the vehicle before and after collision, and the ΔE is the change of the kinetic energy of the vehicle before and after collision.

According to the conservation of kinetic energy, the energy loss of the vehicle in the x-axis direction is as follows:

$$\Delta E_X = E_{Xk0} - E_{Xk} = \left(\frac{1}{2} m_1 v_{A0X}^2 + \frac{1}{2} m_2 v_{B0X}^2 \right) - \left(\frac{1}{2} m_1 v_{A1X}^2 + \frac{1}{2} m_2 v_{B1X}^2 \right)$$

$$= \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \cos \alpha_{A0} - v_{B0} \cos \beta_{B0})^2 \quad (20)$$

where the E_{Xk0} and the E_{Xk} are the kinetic energy of the vehicle before and after the x-axis collision.

Similarly, the energy loss of the vehicle in the z-axis direction can be obtained as follows:

$$\begin{aligned} \Delta E_Z &= E_{Zk0} - E_{Zk} = \left(\frac{1}{2} m_1 v_{A0Z}^2 + \frac{1}{2} m_2 v_{B0Z}^2 \right) - \left(\frac{1}{2} m_1 v_{A1Z}^2 + \frac{1}{2} m_2 v_{B1Z}^2 \right) \\ &= \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \sin \alpha_{A0} - v_{B0} \sin \beta_{B0})^2 \end{aligned} \quad (21)$$

where the E_{Zk0} and E_{Zk} are the kinetic energy of the vehicle before and after the z-axis collision.

Energy absorbed due to vehicle collision E_A and E_B , the relationship is:

$$\frac{E_A}{E_B} = \left(\frac{m_A}{m_B} \right)^2 \quad (22)$$

Namely:

$$E_A = \frac{\Delta E}{1 + \left(\frac{m_B}{m_A} \right)^2} = \Delta E \frac{m_A^2}{m_A^2 + m_B^2} \quad (23)$$

$$E_B = \frac{\Delta E}{1 + \left(\frac{m_A}{m_B} \right)^2} = \Delta E \frac{m_B^2}{m_A^2 + m_B^2} \quad (24)$$

Thus, the change of kinetic energy of car A in the x-axis and z-axis directions can be obtained ΔE_{AX} and ΔE_{AZ} . The change of kinetic energy of car B in x-axis direction and z-axis direction can be obtained ΔE_{BX} and ΔE_{BZ} . The relationship are as follows:

$$\Delta E_{AX} = \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \cos \alpha_{A0} - v_{B0} \cos \beta_{B0})^2 \frac{m_A^2}{(m_A^2 + m_B^2)} \quad (25)$$

$$\Delta E_{AZ} = \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \sin \alpha_{A0} - v_{B0} \sin \beta_{B0})^2 \frac{m_A^2}{(m_A^2 + m_B^2)} \quad (26)$$

$$\Delta E_{BX} = \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \cos \alpha_{A0} - v_{B0} \cos \beta_{B0})^2 \frac{m_B^2}{(m_A^2 + m_B^2)} \quad (27)$$

$$\Delta E_{BZ} = \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \sin \alpha_{A0} - v_{B0} \sin \beta_{B0})^2 \frac{m_B^2}{(m_A^2 + m_B^2)} \quad (28)$$

The deformation energy of vehicle collision is analyzed from the work done by collision force. In the process of vehicle collision, it can be approximately regarded as incomplete elastic deformation. All the work done by collision force F is converted into deformation energy.

$$\Delta E_{def} = W \tag{29}$$

where the W is the work of the external force on the vehicle.

The relationship between collision force and deformation is as follows:

$$F = kx \tag{30}$$

where the k is the elastic recovery coefficient, which is related to the collision material and mass.

$$k = c \frac{m_A m_B}{m_A + m_B} \tag{31}$$

where the c is the spring stiffness coefficient of the front of the vehicle body per unit mass [14]. For the car of 1970s, it is almost a fixed value that has nothing to do with the car type, it's about 41.0 g/m (401.8 m/s²/m).

Since all the work done by the external force on the vehicle is converted into deformation energy, assuming that the deformation amount of the vehicle is s , the work done by the combined external force on the vehicle deformation process is:

$$W = \int_0^s F dx \tag{32}$$

It can be concluded that the work of car A bearing force in x -axis direction in case of collision on steep road section is as follows:

$$W_{AX} = \int_0^{x_{AX}} F_{BX} dx = \int_0^{x_{AX}} G_B \cot \beta_B dx = m_B g x_{AX} \cot \beta_B \tag{33}$$

where the x_{AX} is the deformation of car A in the x -axis direction. The m_B is the mass of car B, and g is the local gravity acceleration.

Similarly, the work done by the force of car A in the z -axis direction can be obtained as follows:

$$W_{AZ} = \int_0^{z_{AZ}} F_{BZ} dz = \int_0^{z_{AZ}} G_B dz = m_B g z_{AZ} \tag{34}$$

where the W_{AZ} is the deformation of car A in the z -axis direction.

The work of the force borne by car B in the x-axis direction is:

$$W_{BX} = \int_0^{x_{BX}} F_{AX} dx = \int_0^{x_{BX}} G_A \cot \alpha_A dx = m_A g x_{BX} \cot \alpha_A \quad (35)$$

where the x_{BX} is the deformation of car B in the x-axis direction, and the m_A is the mass of car A.

The work of the force borne by vehicle B in the z-axis direction is:

$$W_{BZ} = \int_0^{z_{BZ}} F_{AZ} dz = \int_0^{z_{BZ}} G_A dz = m_A g z_{BZ} \quad (36)$$

where the z_{BZ} is the deformation of vehicle B in z-axis direction.

From the conservation of energy in the process of vehicle collision, it can be concluded that the value of the change of kinetic energy of car A in the x-axis direction into the deformation energy is equal to the value of the work done by the external force into the deformation energy, namely:

$$\Delta E_{AX} = W_{AX}$$

The deformation of car a in the x-axis direction can be obtained as follows:

$$x_{AX} = \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \cos \alpha_{A0} - v_{B0} \cos \beta_{B0})^2 \frac{m_A^2}{(m_A^2 + m_B^2)} \frac{\tan \beta_B}{m_B g} \quad (37)$$

Similarly, the deformation of car A in the z-axis direction can be obtained as follows:

$$x_{AZ} = \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \sin \alpha_{A0} - v_{B0} \sin \beta_{B0})^2 \frac{m_A^2}{(m_A^2 + m_B^2)} \frac{1}{m_B g} \quad (38)$$

The deformation of car B in the x-axis direction is:

$$x_{BX} = \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \cos \alpha_{A0} - v_{B0} \cos \beta_{B0})^2 \frac{m_B^2}{(m_A^2 + m_B^2)} \frac{\tan \alpha_A}{m_A g} \quad (39)$$

The deformation of car B in the z-axis direction is:

$$z_{BZ} = \frac{1}{2} \frac{m_A m_B}{m_A + m_B} (1 - \varepsilon^2) (v_{A0} \sin \alpha_{A0} - v_{B0} \sin \beta_{B0})^2 \frac{m_B^2}{(m_A^2 + m_B^2)} \frac{1}{m_A g} \quad (40)$$

According to the above analysis, the deformation of vehicle collision can be obtained by on-site measurement in case of collision accident on the steep slope of

low-grade highway. Because it is assumed that the vehicle does not move after collision, that is, the collision angle before and after collision is unchanged, the relevant angle of collision speed can be obtained, and then the relevant known parameters can be substituted to obtain the vehicle speed before collision, which is an important indicator for traffic police to determine the accident. Due to the limited conditions, it is difficult to test the above mathematical model seriously, so the software widely used in traffic accident simulation is selected for relevant simulation analysis.

4 Pc-Crash Simulation and Analysis

4.1 Brief Introduction of Simulation Software

Pc-Crash is a traffic accident reconstruction software. Based on the dynamic analysis of the causes of traffic accidents, Austrian scholars have analyzed the on-site information of traffic accidents for many years to deduce the process of vehicle collision from the post collision position of the accident vehicle, and analyze the movement state of the vehicle before the accident, so as to analyze the causes of the accident. The software has been handed in Police are widely used in traffic accident analysis, which is of great significance for quick accident determination.

According to the characteristics of traffic accident scene, such as vehicle position, vehicle deformation and road condition after collision, the simulation software uses the basic principle of energy conservation to reverse the accident process. The external environment of the road section can be created by inputting file information, and the collision process can be analyzed more intuitively under the condition of 3D view. Through the function key, the slope, turning radius and cross slope of the road section can be adjusted, the road slope information can be modified to meet the actual situation of vehicle traffic, and the vehicle operation can be further changed by adjusting the adhesion coefficient of the road. The vehicle model can be input through the database provided by the system software, or the model can be built by itself to input relevant information. Both of them can modify the basic attributes such as the vehicle's dynamic performance, vehicle's three-dimensional data, tire configuration, vehicle's braking performance through the toolbar, so as to be more close to the actual situation of the collision vehicle. Through the system's own database, the driver can be embedded in the simulation model to get close to the actual situation of vehicle driving. Its reconstruction work is to study the vehicle speed track before and after the accident through the statistical regression analysis and prediction, accident scene simulation, collision track analysis and other methods, and finally output the calculation results or simulated operation in the form of screen display, image and data accident process.

4.2 Simulation and Analysis

In this paper, by analyzing the characteristics of steep road sections and driving, a simulation model is built, which accords with the characteristics of sharp curve and steep road sections. By adjusting the road surface information and road slope, the model road condition information is close to the actual road information. The car with strong climbing ability is selected as the vehicle, and the same type of vehicle is selected as the simulation vehicle, and the same running state is set. Through the dynamic simulation, the collision deformation of the vehicle is obtained. Through the amount of vehicle collision deformation under different conditions, the amount of collision deformation caused by different speeds is compared and analyzed.

In this paper, the research road section is the steep slope section of low-grade highway, and the vehicle needs strong climbing ability. The simulation vehicle chooses Ford-ecosport when building the vehicle collision model. The whole vehicle mass is 1350 kg. The length is 4.23 m. The width is 1.74. Assuming that the vehicle center of gravity is located in the vehicle center. The maximum engine horsepower is 102hp, and the motor vehicle engine is 1.5 L double independent camshaft variable timing engine. According to the analysis in Chap. 2, the slope of the road is set as 0.06. The initial speed is set as 50 km/h. The radius of the vertical curve is designed as 700 m according to the general value. The friction coefficient of the road is set as 0.8 according to the dry asphalt pavement. The lanes are two-way two lanes., The vehicle generally drives along the road center line when the vehicle is driving in the steep slope sharp turn road section. In case of rain, snow and other weather, the vehicle chooses to drive along the road center line. At the same time, the line of sight is limited by the ramp, so it is difficult to observe the road information in front. The vehicle is easy to have a frontal collision in the slope top area. In the high subgrade road section, it is also easy to have traffic accidents such as rollover and side collision, and cause great property loss and heat source casualties. The specific 3D design model of road section and vehicle is shown in Fig. 5.

The vehicle collides at the top of the steep slope section, and the collision point and deformation are shown in Fig. 6.

Fig. 5 Driving model of vehicle on steep slope of low-grade highway

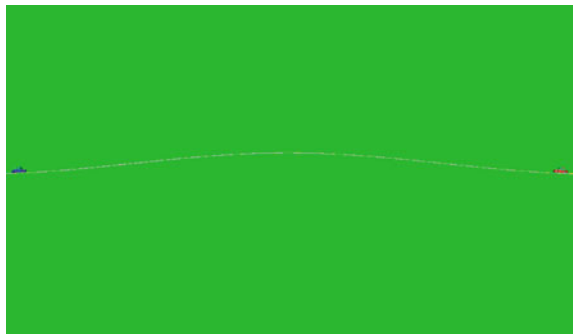
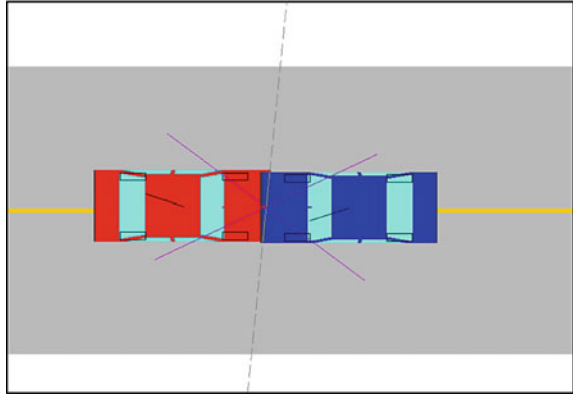


Fig. 6 Vehicle collision on steep slope of low-grade highway.



The dotted line in the figure is the vehicle collision plane. The solid line is the vehicle collision direction. The vehicle deformation is 8 cm when the vehicle speed is 40 km/h. The vehicle deformation is 23 cm when the vehicle speed is 50 km/h. The vehicle deformation is 24 cm when the vehicle speed is 60 km/h. The vehicle deformation is 40 cm when the vehicle speed is 70 km/h. The vehicle deformation is 30 cm when the vehicle speed is 80 km/h. The analysis shows that when the vehicle speed is below 60 km/h, the deformation of the vehicle increases with the increase of the speed. The deformation of the vehicle will change suddenly when the vehicle speed reaches more than 60 km/h. At this time, the vehicle is seriously damaged and the deformation of the vehicle is large, which is easy to cause casualties.

5 Conclusion and Prospect

5.1 Conclusion

In this paper, the driving distance of vehicles on the steep slope is mainly analyzed by analyzing the unsafe factors existing in the driving of vehicles on the steep slope of the low-grade highway. The vehicle collision problems caused by the driving sight distance and the special meteorological conditions are analyzed. The vehicle collision process caused by the sight distance is quantitatively analyzed, which is a low-grade highway. The safety protection measures of vehicle collision provide relevant reference.

Based on the analysis of the sight distance of the steep road section, the dynamic analysis of vehicle collision is carried out. The loss of kinetic energy in the process of vehicle collision is all converted into deformation energy, and then the force analysis of the process of vehicle collision is carried out. The combined external force work of the vehicle is all converted into deformation energy, and the loss of kinetic energy and combined external force in the process of vehicle collision can be obtained

by energy conservation. Based on the conservation of quantity, the mathematical model of vehicle deformation and speed can be obtained, which provides the research direction for the analysis of the relationship between vehicle deformation and speed in steep road section.

Based on the Pc-Crash traffic accident simulation software, this paper analyzes the relevant parameters of low-grade highway steep slope section and the relevant parameters of the driving vehicles in this section, constructs the vehicle collision model of the vehicles in the steep slope section, analyzes the different simulation results, and concludes that when the vehicle speed reaches 60 km/h or above, the vehicle deformation will occur sudden change, at this time, the vehicle damage is serious, the vehicle deformation is large and easy to build Conclusion of casualties.

5.2 Prospect

In this paper, only the sight distance of vehicle collision in the steep road section of low-grade highway and the driving characteristics of vehicle under special weather conditions are analyzed. Subsequently, the traffic sign and road alignment of vehicle in the driving process of this type of road section can be analyzed, and the impact of different road sections on the driver can be quantitatively analyzed, so as to provide a basis for the subsequent safety driving of this road section on the basis of analysis. In addition, it can also analyze the relevant situation of vehicle collision in the sharp curve section and the steep and accelerated curve section, explore the risk factors that affect the safe driving of vehicles in the special section of low-grade highway, and provide theoretical basis for the follow-up to provide reasonable safety measures.

In this paper, only Pc-Crash software is used to build the relevant models of the same kind of vehicle and the same kind of vehicle's frontal collision, and the simulation analysis is carried out. In addition, the relevant simulation analysis can be carried out for different vehicle's collision problems in different road types and different vehicle speeds, so as to verify the accuracy of the mathematical model from multiple perspectives. Due to the complexity of vehicle collision process, this paper only considers some of them, ignoring other factors that have little impact on collision. In order to grasp the relationship between deformation and vehicle speed in the process of vehicle collision more accurately, other factors can be considered to improve the accuracy of collision deformation mathematical model.

References

1. Li M (2016) Study on roadside safety risk identification and improvement measures of Mountainous Expresswa. Chang'an University (in Chinese)
2. Zhang C, Ma R (2010) Traffic accident analysis and frequent section identification of Mountainous Expressway. J Chang'an Univ (Natural Science Edition) 30(06):76–80 (in Chinese)

3. Road Science Research Institute of the Ministry of Transport (2013) China road traffic safety blue book. People's Traffic Press, Beijing, pp 29–30 (in Chinese)
4. Yutao F, Quancai W, Jun L (2020) Disease causes and treatment techniques of high embankment on rich water steep slope of mountain expressway. *Disaster Sc* 35(01):90–93 (in Chinese)
5. Tao X, Yuming J, Jian L, Fei L (2020) Study on driving safety of mountain highway based on reliability method. *J Wuhan Univ Technol (Traffic Science and Engineering Edition)* 44(01):120–125 (in Chinese)
6. Dissanayake S, Roy U (2014) Crash severity analysis of single vehicle run-off-road crashes. *J Transp Technol* 4(1):1–10
7. Bambach MR, Mitchell RJ, Grzebieta RH (2013) The protective effect of roadside barriers for motorcyclists. *Traffic Injury Prevent* 14(7)
8. Carlos R, Filipe M, João LC (2015) Detecting unforgiving roadside contributors through the severity analysis of ran-off-road crashes. *Accident Anal Prevent* 80
9. Zuoren Y, Yuren C et al (2012) Road engineering. People's Communications Press, Beijing (in Chinese)
10. Wen X (2017) Study on search range of dynamic vision of grassland road drivers. Inner Mongolia Agricultural University (in Chinese)
11. Guoliang Z, Xubin D, Liping D (2012) Design of intersection safety sight distance and intersection under viaduct. *Urban Road Bridge Flood Control* 10:3–5 (in Chinese)
12. Zhaoxiang Y (2002) Course of traffic accident mechanics appraisal. Guangxi University of Science and Technology, Nanning (in Chinese)
13. Bao S (2009) Research on vehicle crash simulation and reconstruction. Southwest Jiaotong University (in Chinese)
14. Xu H (2013) Automobile accident engineering. Jilin University (in Chinese)

Research on the Influence of Freeway Truck Mixing Rate on Traffic State



Fengchun Han, Na Li, Dan Zhao, and Yumeng Zhang

Abstract As the high-speed freight passage between cities, the freeway plays an important role in promoting the economic development along. Freight occupies a large proportion of the traffic on freeway. The characteristics of trucks, such as big size, poor flexibility, slow speed, would make a certain impact on the traffic state. In this paper, based on the traffic survey data of typical sections in mountainous areas, regression was used to study the relationship between truck mixing rate and traffic flow parameters. By clustering analysis of traffic flow data, the distance-based K-means algorithm was adopted to demarcate the freeway traffic state and determine the threshold of the influence of truck mixing rate on the traffic state. The results show that with the premise of a certain speed and density, when the truck mixing rate is less than 0.35, a good traffic state is identified; when the rate is higher than 0.35, a greater impact on the traffic flow is spotted, leading to a high possibility of traffic accidents. Hence, it is concluded that freight traffic should be reasonably controlled to enter the freeway.

Keywords Freeway in mountainous areas · Truck mixing rate · Regression analysis · Cluster analysis · Traffic state

F. Han · D. Zhao (✉) · Y. Zhang
School of Traffic Management, People's Public Security University of China, ST Huangyi,
Beijing, China
e-mail: zhaodan@ppsuc.edu.cn

F. Han
e-mail: hfc1966@163.com

Y. Zhang
e-mail: m13520312497@163.com

N. Li
Fujian Police College, No. 59 Shoushan Road, Cangshan District, Fuzhou, Fujian, China
e-mail: fjplina@126.com

1 Introduction

According to the Statistical Communiqué [1], until the end of 2019, the total length of highways in our country reached 5.0125 million km, including 149,600 km of freeways, which overtook America to become the one who has the longest highways in the world. With the rapid development of China's economy, freeways undertake a large number of passenger and freight traffics. The demand for freight transportation grows every year. When the proportion of trucks reaches a certain number, the efficiency of freeway traffic flow and the level of service will be sacrificed, and the traffic flow will be more dispersed, which often results traffic accidents. Therefore, it is of great practical significance to research the influence of the truck mixing on traffic state of freeway.

Freight transportation is earlier studied in foreign countries, and most of the studies focus on safety and mobility bottleneck theory. Prigogine and Herman [2] first proposed "Moving Bottleneck", meaning the lower speed of large vehicles will cause several or even a large number of vehicles to slow down and keep following as a "clustered" fleet, which reduces the traffic capacity and increases vehicle delays; In order to explore the problem of traffic congestion, Lozano et al. [3] and Montazeri-Gh and Fotouhi [4] classified the traffic state by the K-means clustering method based on the actual survey of traffic flow parameter; Based on the principle of big data analysis, Li and Porikli [5] and Porikli and Li [6] constructed a traffic state classification method by using Gaussian mixture hidden Markov model; Leclercq et al. [7] introduced the bottleneck diagram to analyze the traffic state on the adjacent lanes of the moving bottleneck, linking the traffic state on the upstream and downstream of the moving bottleneck, unifying the models proposed by different scholars, but it lacks the support of actual survey data; Daganzo and Laval [8] of the United States tested the condition of the freeway, while they first described the characteristics of a single moving bottleneck. Under the premise of counting a large amount of data, the formula was proposed which is about the change of downstream dissipation flow with the speed of the moving bottleneck when it occurs; Based on the principle of intelligent data classification, Shin et al. [9] used measured data to classify the traffic operation status and verified the accuracy of the classification. The first domestic research of the trucks mixed rate on freeway was done in 2005. Gao et al. [10] studied the impact of trucks on the safety of freeway operations based on statistical data of typical freeway traffic accidents. The corresponding improvement measures from Traffic management, engineering and other perspectives were proposed; Zhuang and Lv [11] conducted a field survey of traffic condition in Shenzhen, obtained the parameters of speed and density, fitted different traffic flow models, and found that Greenshields has the best linear relationship fitting effect; Xu et al. [12] took the Beijing-Tianjin-Tangshan freeway as an example to analyze the influence of the time distribution of tracks on the speed and the speed difference of traffic flow. Based on the variable speed limit of the proportion of large cars, the process and speed limit was determined and the impact of the large cars mixing rate on the speed was analyzed; Liu [13] studied on the moving bottleneck on freeway and

simulated different traffic volume and truck mixing rate by Vissim to find the rules of the changing of traffic flow parameters with the moving bottleneck; Based on the analyzation of the influence of trucks on freeway traffic flow, Wang [14] simulated the volume and the speed of the traffic flow on freeway by Vissim to obtain the general conditions for setting up the lanes of truck; Liang et al. [15] analyzed the relationship between the truck mixing rate and the stability of traffic flow on freeway, and treated the standard deviation of the following distance, rate of the vehicle lane change, the standard deviation of speed, and the average delay as the evaluation indexes of the traffic flow stability. The truck mixing rate was researched, the vehicle parameters and variable parameters were designed, and the traffic flow conditions in different traffic volume, number of one-way lanes, and longitudinal slope were simulated to divide the stability of the traffic flow into four levels.

In order to reveal the influence of trucks on the traffic state in different mixing rates, take the measured traffic data of typical mountain section of Jinji freeway as an example, mathematical statistics, clustering algorithms and computer programming methods were used to study the relationship between truck mixing rate and traffic state. The threshold of the impact of the truck mixing rate on the traffic state was obtained by cluster analysis to reasonably control the proportion of trucks, and improve the efficiency and the traffic safety on freeway.

2 Traffic Survey and Data Acquisition

2.1 Traffic Survey

Based on the on-site traffic data of the typical section of Jinji Freeway, the traffic volume with different traffic compositions and location speed in a certain period of time, taking account the truck mixing rate, were counted. Thus, the influence rule of the truck mixing rate on the traffic state was obtained.

- (1) Survey time: For meeting the purpose of the survey, the parameter of traffic volume and location speed was investigated for 9 consecutive hours, from 8:00 to 17:00 on December 30, 2019.
- (2) Survey location: The research object is the basic section of Jinji Freeway (as shown in Fig. 1), which has straight and flat alignment, with four lanes in both directions and a maximum speed limit of 100 km/h. In order to investigate the location speed on the basic section, the average speed of the traffic flow is greater than 70 km/h, so the length of the survey section set to be 80 m.
- (3) Survey method: Collect the traffic flow parameters in the section from surveillance video by manual counting.
- (4) Survey content: Investigate the traffic volume and location speed on the section for 5 min, and obtain the truck mixing rate.
- (5) Survey sample size: In order to meet the statistical requirements and ensure the accuracy of the data, the survey of traffic flow parameter was conducted. Set

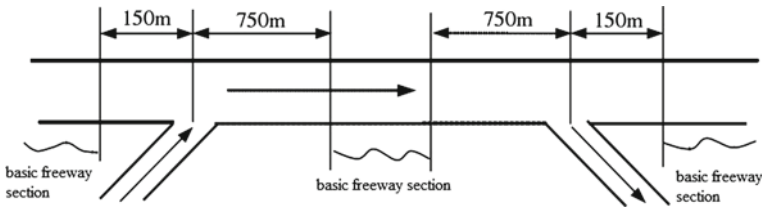


Fig. 1 The basic freeway section

the confidence to 95%, the sample standard deviation to 8 km/h, the allowable error to 2 km/h, and the minimum sample size to 62.

- (6) Survey requirements: In normal working days without bad weather, holidays and large-scale events, etc., taking 5 min as the counting interval, recording the traffic parameters of different vehicle models in one direction, 108 sets of survey data were obtained, which can meet the minimum sample size requirements.

2.2 Data Acquisition

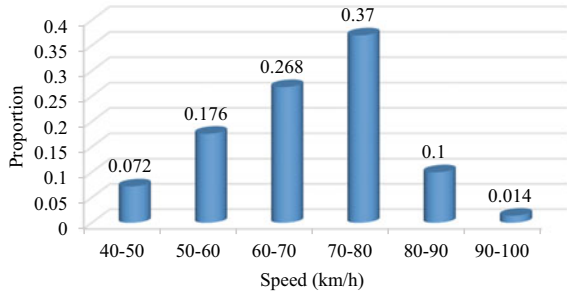
- (1) Obtain the traffic volume of 5 min. The 9 h surveillance video data of the observation location was played frame by frame, and the equivalent traffic volume of different models would be counted.
- (2) Calculate the location speed. The start and the end of the survey section that length of 80 m were determined, the moment of each vehicles arrive the start section and end section to calculate the time of each vehicle passing the survey section were recorded, and the location speed was obtained by dividing the distance by the time.
- (3) Calculate the truck mixing rate. The truck mixing rate refers to the ratio of truck to the all of vehicles in a certain road space and a certain period of time. The truck mixing rate reflects the change characteristics of the traffic state caused by trucks over time within a certain period. The truck mixing rate has no unit, and the value is between 0 and 1.

3 Analysis of Truck Mixing Characteristics

3.1 Truck Speed Characteristics

Compared with cars, the speed of trucks is relatively lower. When the proportion of trucks reaches a certain value, the speed difference between vehicles widens, and the traffic flow on the freeway is in an unstable state. Based on the location speed statistics of each vehicle types in the survey section, the distribution characteristics

Fig. 2 Distribution of truck speeds on the observed freeway section



of truck speed were obtained, as shown in Fig. 2. The speed of trucks basically conformed to the normal distribution in Fig. 2. The highest proportion reached 37% of truck speeds at 70–80 km/h, followed by 60–70 km/h at 26.8%, and few of speeds were above 90 km/h and below 50 km.

3.2 Traffic Characteristics of Truck

To make the chart of truck proportion for the corresponding period, as shown in Fig. 3, the truck volume per hour was counted. The ratio of trucks in the survey section was about 0.4, and the ratio of trucks from 13:00 to 14:00 in the afternoon reached a maximum as 0.46. Therefore, the impact of trucks on traffic flow operation cannot be ignored.

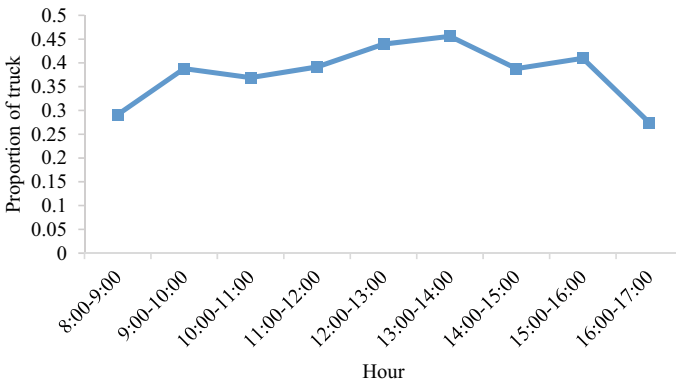


Fig. 3 Proportion of truck to the total volume of vehicle on observed freeway section

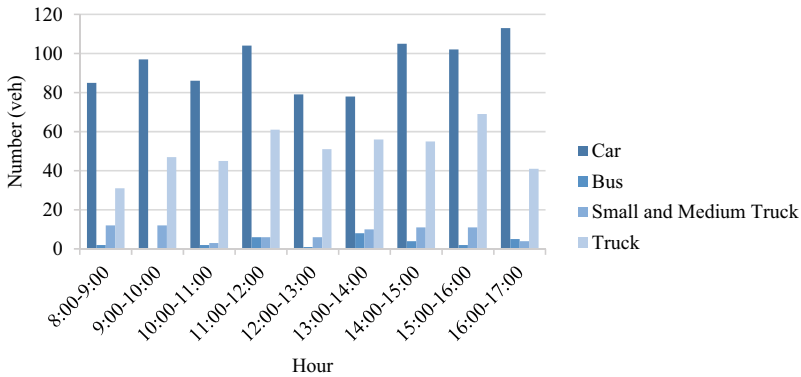


Fig. 4 Distribution of each vehicle types on the observed freeway section

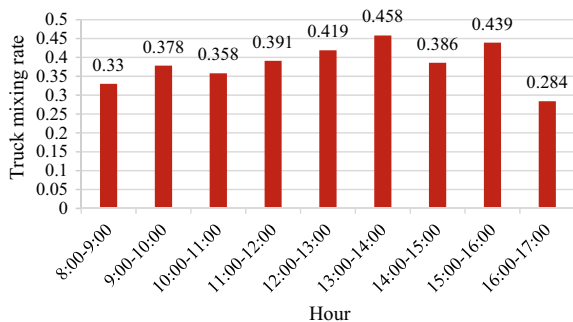
3.3 Distribution Characteristics of Truck Models

The observed section belongs to the inter-city freeway which had high proportion of trucks. The vehicles in the observed section are divided into cars, small and medium trucks, buses, and trucks. The statistical results were shown in Fig. 4. The proportion of small cars is the highest as 57%, small and medium trucks account for 4%, buses accounts for 1%, and trucks is the second highest, up to about 38%.

3.4 Characteristics of Truck Mixing Rate

Counted the 5 min truck volume and total volume on the observed section, and calculated the change of truck mixing rate with time, as shown in Fig. 5. The average truck mixing rate is 0.38, the highest is 0.45, and the lowest is 0.284.

Fig. 5 Distribution of truck mixing rate over time on observed freeway section



4 The Relationship Between Truck Mixing Rate and Traffic State

4.1 Relationship Between Truck Mixing Rate and Speed

Trucks had large volume, poor flexibility, and occupy large driving space. With the increase of truck mixing rate, the interference between vehicles becomes more serious, and the overall speed of freeway traffic flow decreases. According to the statistical data, a scatterplot of truck mixing rate and running speed is shown in Fig. 6.

The speed of freeway traffic flow decreases with the increase of the truck mixing rate. When the truck mixing rate is between 0 and 0.6, as the truck mixing rate increases, the downward trend of traffic flow speed becomes more obvious; when the truck mixing rate reaches 0.6–0.8, the traffic flow speed decreases slowly and gradually stabilizes. Linear and quadratic polynomials are used to fit the truck mixing rate and speed scatterplot respectively. The expressions and coefficient of determinations are shown in Table 1.

It can be seen from Table 1 that the coefficient of determination of the quadratic polynomial model is large, which is selected to describe the relationship between the truck mixing rate and the operational speed of the traffic in the survey section. The expression is:

$$v = 40.516r^2 - 68.349r + 110.29 \tag{1}$$

Fig. 6 Quadratic polynomial fitting curve of truck mixing rate and speed

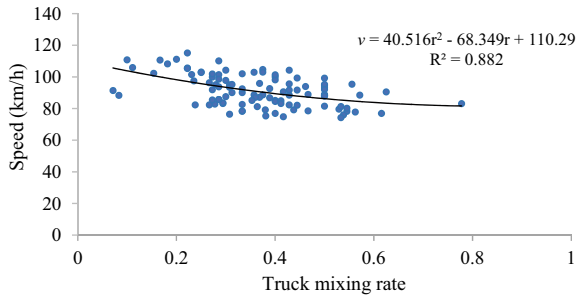


Table 1 Regression results of the relationship between truck mixing rate and speed

Model type	Expression of the model	Coefficient of determination (R ²)
Linear model	$v = -37.699r + 105.13$	0.729
Quadratic polynomial model	$v = 40.516r^2 - 68.349r + 110.29$	0.882

where r is the truck mixing rate;
 v is traffic flow speed, in km/h.

4.2 Relationship Between Truck Mixing Rate and Volume

When the truck mixing rate on freeway is low, the interference of trucks to other vehicles is less, and the passing traffic volume per unit time increases as the truck mixing rate grows; when the flow of freeway rises to a certain value, the impact of trucks on traffic flow becomes greater in a way that the flow rate per unit time decreases as the truck mixing rate increases. According to the statistical data, the scatter diagram of truck mixing rate and volume is shown in Fig. 7.

The flow rate of the freeway per unit time increases first and then decreases as the truck mixing rate grows. When the truck mixing rate is less than 0.5, as the truck mixing rate increases, the rate climbs significantly; when the truck mixing rate reaches 0.5–0.6, the rate tends to be larger; as the truck mixing rate is greater than 0.6, the flow rate slowly declines. Both linear polynomials and quadratic polynomials are selected to fit the relationship between truck mixing rate and volume. The results indicate that when using the quadratic polynomial model to fit the scatter plot of truck mixing rate and volume, the coefficient of determination is 0.821, which is better than the linear. The expression is:

$$q = -38.923r^2 + 44.397r + 16.543 \tag{2}$$

where r is truck mixing rate;
 q is the volume pcu/h.

By analyzing the impact of truck mixing rate on speed and volume, the following conclusions are obtained:

- (1) When the truck mixing rate on freeways is less than 0.5, as the truck mixing rate increases, the speed drops significantly. When the truck mixing rate is greater than 0.6, the speed decreases slowly.

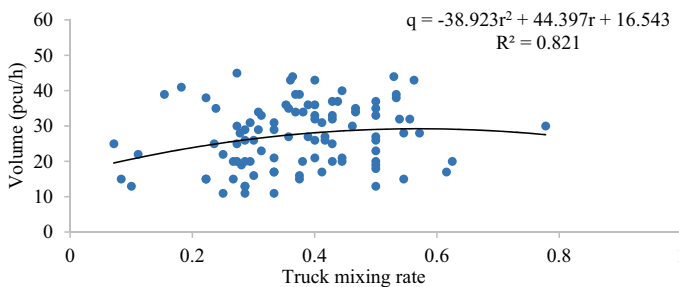


Fig. 7 Scatterplot of the relationship between the truck mixing rate and the volume

- (2) The volume of the observed freeway section changes along—increases first and then decreased with the upward truck mixing rate, and reaches the maximum when the truck mixing rate reaches 0.6.

Above all, considering the influence of freeway truck mixing rate on traffic flow parameters, when the truck mixing rate reaches 0.6, it would have a greater impact on the freeway traffic state.

4.3 Cluster Analysis of Traffic Data

The clustering algorithm [16] can divide the survey data into several clusters, so that the data with higher similarity were grouped together. The intensity of data, the overall distribution pattern, and the certain relationships between the data attributes were obtained by clustering. When using a clustering algorithm, the gap between data needed to be calculated and expressed as a distance. This article used Euclidean distance [17] calculation method which is the most commonly used distance calculation method:

$$d(i, j) = \sqrt{|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2} \tag{3}$$

where i, j respectively represents two data in the data set, they have p attributes.

K-means algorithm [18] is one of the classic clustering algorithms. Because of its high efficiency, it is widely used in data mining. In this paper, the K-means algorithm was used to classify the traffic state of the observed freeway section by cluster analysis of the traffic data.

4.3.1 Clustering Algorithm

As evaluation indicators, traffic volume, speed, and density were used to divide the freeway operating status, determining the data set n and the number of clusters k to be divided. In view of the research purpose, n was 108 pairs of data (speed, volume, density) obtained from the survey on Jinji freeway, and k was defined into three categories: the clustering center no. 1 meant traffic state is very good; the clustering center no. 2 meant the traffic state is good; the clustering center no. 3 meant the traffic state is bad.

The specific calculation process was as follows:

- (1) Selected three ($v_1, q_1; v_2, q_2; v_3, q_3$) from 108 pairs of data sets to be the initial clustering center;
- (2) Used Eq. (3) to calculate the distance of each data. For example, the next data was (v_4, q_4), then calculate which d was closest to the initial three pairs of data;

Table 2 Ranges of traffic volume and speed in three operating status

Cluster	Operating status	Volume (pcu/h)	Speed (km/h)
1	Very good	(9, 22)	(85.4, 115.1)
2	Good	(23, 33)	(74.8, 102.9)
3	General	(34, 56)	(74.2, 95.9)

- (3) If the calculated d was closest to $(v1, q1)$, then $(v4, q4)$ and $(v1, q1)$ were classified into one category;
- (4) Update clustering center. The average value of the two pairs of data classified into one category was taken as the new clustering center of this category, and the other two center clusters were unchanged;
- (5) By analogy, until all of data are correctly classified, the clustering center would be constant.

4.3.2 Calculation Results

In this paper, C++ was used for programming, and the data of 108 pairs of traffic volume and speed at intervals of 5 min were used to calculate the clustering results. The obtained data was therefore divided into three clusters, which were used to determine the three traffic operating status ranges corresponding to the traffic volume and speed, as shown in Table 2.

3 clusters were obtained by the above data analysis:

The first cluster contains 40 cases. The characteristics as follow, the time interval was mostly concentrated at 8:00–11:00 a.m. During this period, the quantity of freeway trucks and the total traffic volume were small, the speed and road capacity were high, and the traffic status was very good.

The second cluster contains 36 cases. The characteristics as follow, in the cluster time interval, the traffic volume, the quantity of trucks and the speed was medium, and the traffic status was good.

The third cluster contains 32 cases, and the characteristics as follow, the time interval was concentrated at 11:00–12:00 a.m. and 14:00–16:00 p.m. In this period, the number of trucks on freeway was relatively large, at 14:00–15:00 p.m. the truck mixing rate reached the maximum value in a day. The impact on the traffic flow became significant, the average speed of the vehicle flow lowered down, the road capacity decreased, and the traffic status was general.

4.3.3 Data Analysis

The operation status has three levels: very good, good, and general, the time interval of each cluster and the corresponding truck mixing rate were counted.

The first cluster time interval was mostly concentrated between 8:00–9:00 and 10:00–11:00. From the perspective of the truck mixing rate in each time period, there

Table 3 The truck mixing rate corresponding to the time interval in the third cluster

Time interval	Truck mixing rate	Time interval	Truck mixing rate
9:15–9:20	0.5	9:20–9:25	0.4375
9:25–9:30	0.411765	9:45–9:50	0.428571
10:40–10:45	0.444444	10:55–11:00	0.375
11:15–11:20	0.4	11:20–11:25	0.38
11:30–11:35	0.388889	11:45–11:50	0.363636
12:25–12:30	0.529412	12:35–12:40	0.368421
12:50–12:55	0.571429	13:00–13:05	0.428517
13:10–13:15	0.5	13:40–13:45	0.5
13:45–13:50	0.352941	13:50–13:55	0.615385
14:00–14:05	0.28	14:10–14:15	0.411765
14:35–14:40	0.533333	14:45–14:50	0.388889
14:55–15:00	0.466667	15:00–15:05	0.368421
15:10–15:15	0.545455	15:15–15:20	0.533333
15:20–15:25	0.36	15:25–15:30	0.5625
15:30–15:35	0.3	15:45–15:50	0.4
16:25–16:30	0.238095	16:55–17:00	0.466667

were 23 cases with a mixing rate less than 0.3, accounting for about 60%; there were 7 cases with mixing rate above 0.4. During this period, there was few vehicles on the freeway, the average speed of the vehicles was higher and the interference between the vehicles was less. Even if the truck mixing rate grew, the operation status was still good.

The truck mixing rate in the second cluster was mostly between 0.22 and 0.35, accounting for about 62% of the total, and there were 10 cases' mixing rate was above 0.4. During this period, the vehicle speed was high and the total traffic volume was low. The change of the truck mixing rate did not caused a significant impact on other vehicles, and the traffic operation was stable.

The truck mixing rate in the third cluster (as shown in Table 3) was more than 0.35, accounting for about 90% of all data. Among them, the maximum value of the truck mixing rate between 13:50 and 13:55 was 0.62, and the traffic state was average.

In summary, on a mountainous freeway with two driving directions and four lanes, the truck mixing rate has a certain impact on the traffic state. When the truck mixing rate is less than 0.35 (corresponding to a volume of 33 pcu/h), the freeway traffic is in good condition; when the truck mixing rate is greater than 0.35, the impact on the traffic flow increases, the traffic state suffers, and the service level compromises.

5 Conclusions

To study the impact of the truck mixing rate on the traffic state, based on the 9 h continuous traffic flow data of the basic section of the Jinji freeway, the regression analysis was used to obtain the functional relationship between the truck mixing rate and the traffic volume or speed. The K-means clustering algorithm was used to divide the freeway traffic state into three categories: very good, good, and general. The results show that the threshold of the impact of freeway truck mixing rate on traffic flow is 0.35, and the corresponding flow is 33 pcu/h. When the truck mixing rate is less than 0.35, the freeway traffic operation state is stable. When the truck mixing rate is greater than 0.35, the impact on the traffic flow increases, the traffic state is general, and the freeway service level declines.

The main contribution of this research is to use the clustering K-means algorithm to divide the traffic state on the basic freeway section, and to use the relationship between the truck mixing rate and the traffic volume or speed to established the law of the influence of truck mixing rate on the traffic state of basic freeway sections, which breaks through the limitations of previous researches using traffic simulation methods. This research is based on actual survey data, and the results lays a theoretical foundation for the in-depth study of the relationship between the truck mixing rate on the mountain freeway and the traffic state.

The conclusion of the study provides a reference for the freeway management department to control the inflow of trucks on mountainous freeways according to the traffic status. In the follow-up study, an in-depth research could be carried out for different regions and different road alignments with prolonged investigation time and larger and structural data to analyze the changes in traffic state caused by the mixing of different trucks.

References

1. Statistical Communique on the 2019 transportation industry development
2. Prigogine L, Herman R (1971) Kinetic theory of vehicular traffic. Elsevier, New York
3. Lozano A, Manfredi G, Nieldu L (2009) An algorithm for the recognition of levels of congestion in road traffic problems. *Math Comput Simul* 79(6):1926–1934
4. Montazeri-Gh M, Fotouhi A (2011) Traffic condition recognition using the k-means clustering method. *Sci Iran* 18(4)
5. Li X, Porikli FM (2004) A hidden Markov model framework for traffic event detection using video features. In: International conference on image processing, vol 5, pp 2901–2904
6. Porikli FM, Li X (2004) Traffic congestion estimation using HMM models without vehicle tracking. In: IEEE intelligent vehicles symposium, pp 188–193
7. Leclercq L, Chanut S, Lesort J-B (2004) Moving bottlenecks in the LWR model: a unified theory. In: The 83rd TRB annual meeting
8. Daganzo C, Laval JA (2005) On the numerical treatment of moving bottlenecks. *Transp Res* 39:31–46
9. Shin D, Aliaga D et al (2014) Urban sensing: using smartphones for transportation mode classification. *Comput Environ Urban Syst* S0198971514000921

10. Gao J, Liu B, Guo Z (2005) Influence of heavy vehicles on traffic safety of freeways. *J Tongji Univ (Nat Sci Ed)* 33(6):763–767
11. Zhuang Y, Lv S (2005) A study of the speed-flow-density relationships on urban roads. *J Shenzhen Univ Sci Technol Ed* 22(4):373–376
12. Xu T, Sun X, He Y et al (2009) Mixed rate of large vehicles in expressway variable speed limit. *Traffic Inf Saf* 2(27):70–72
13. Liu J (2011) Analysis of the impact of expressway bottleneck effect on traffic flow. *Sci Technol Innov Appl* (004):13–14
14. Wang W (2012) Traffic condition research and case study of freeway truck only lane. Master's thesis, Beijing Jiaotong University
15. Liang G, Cheng G, Wang C et al (2014) Relationship between mixing rate of freeway large vehicles and traffic flow stability. *J Chang'an Univ (Nat Sci Ed)* 07
16. Sun JG, Liu J, Zhao LY (2008) Clustering algorithms research. *J Softw* 19(1):48–61
17. Zhang Z, Cao Z, Li Y (2010) Research based on Euclid distance with weights of K-means algorithm. *J Zhengzhou Univ (Eng Ed)* 31(01):89–92
18. Li H (2014) Improvement of K-means clustering method and its application. Northeast Agricultural University

Experimental Method to Establish Vehicles' Acceleration/Deceleration Speed Parameter Thresholds for Urban Roads in Beijing



Chenjing Zhou, Xinyue Nie, Zihe Lin, Jian Rong, and Jianming Ma

Abstract Traffic microsimulation tools can be used to describe complex traffic system features and reproduce traffic scenarios. Calibrating parameters is the prerequisite of using simulation technology scientifically. According to the properties, parameters are divided into basic parameters and model parameters for more accurate use in this paper. The basic parameters act on the global behavior characteristics of the vehicle and the model parameters act on the specific driving behavior. This research designed two actual urban road experiments to perform engineering tests on the basic parameters, including the relation between the interval of acceleration values and the driving experience, and the characteristics of acceleration and speed. This paper is proposed an experimental method that can analyze the acceptable threshold of acceleration, which divided acceleration into three phases: unadaptable and unacceptable phase, unadaptable and acceptable phase, adaptable and acceptable phase. The study also proposed the value interval for basic parameters and took the data envelope area which focused on the data fitting line as the form of

C. Zhou (✉) · X. Nie

Beijing University of Civil Engineering and Architecture, 1 Zhanlan Road, Xicheng District, Beijing 100044, People's Republic of China
e-mail: 13439099781@139.com

X. Nie

e-mail: 0963nie@163.com

Z. Lin

China Highway Engineering Consulting Corporation, 17 Changyungong, Haidian District, Beijing 100089, People's Republic of China
e-mail: linsanitiest@163.com

J. Rong

Beijing Key Laboratory of Traffic Engineering, Beijing Engineering Research Center of Urban Transport Operation Guarantee, Beijing University of Technology, 100 Pingleyuan, Chaoyang District, Beijing 100124, People's Republic of China
e-mail: jrong@bjut.edu.cn

J. Ma

Texas Department of Transportation, 9500 N. Lake Creek Parkway, Austin, TX 78717, USA
e-mail: jianming.ma@txdot.gov

expression of the relation between speed and acceleration: (1) In starting phase ($V < 7.2$ km/h), the acceleration-speed slope is $k = 0.1832$. (2) In increasing phase ($V = 7.2\text{--}35.4$ km/h), the overall slope of the acceleration-speed in this phase is $k = 0.0256$. (3) In stabilizing phase ($V > 35.4$ km/h), the speed is relatively stable, and the acceleration-speed slope is approximately equal to zero. The study is beneficial to the practical application of simulation work and self-driving automobile.

Keywords Microsimulation · Basic parameters · Parameter calibration · Parameter threshold · Data envelope area

1 Introduction

The realistic traffic scene is a complex system with multiple interacting factors. In addition, various characteristics of the system continue to vary in space and time. It is difficult to describe the system operation characteristics and process using traditional mathematical methods. Traffic microsimulation technology uses the event or time as the driving mechanism to evolve the operation process of the transportation system, and overcomes the shortcomings of traditional mathematics, as it can describe the high-dimensional complex transportation system and make possible a wide range of applications. The core research content in this field is to ensure the authenticity of the microsimulation system.

The adjustment of model parameters and site-specific measurement of parameters to improve a model's predictive ability to reproduce driver behavior and traffic characteristics is known as calibration [1]. Benekohal [2] established a framework for parameter validation and verification of traffic microsimulation models. The simulation validation method described by Hellinga [3] can be divided into three parts: the preliminary analysis preparation, the simulation model initialization correction and the evaluation of the analytical result. Brockfeld et al. [4] used high-precision GPS to collect vehicle operation records and calibrated 10 parameters of follow-up models. The calibration results showed that all models can reproduce the vehicle motion process within a certain error range. Based on the simulation validation by Hellinga [3], Park and Schneeberger [5] proposed a simulation model to calibrate the 9-step calibration process. Azam et al. [6] and Guo et al. [7] described the complete calibration work in traffic microsimulation in the application process of various transportation fields. Gomes et al. [8] and Merritt [9] constructed the simulation calibration process to carry out simulation calibration work on the bottleneck point in the congestion state respectively. Rong et al. [10] used traffic simulation tools to solve traffic capacity and calibrated their simulation models by controlling the speed-flow curve difference between actual and simulated data. Ciuffo et al. [11] summarized the overall process of microsimulation calibration, including identifying performance evaluation indicators, performing model sensitivity analysis, selecting optimization algorithms, determining objective functions, conducting parameter optimization based on optimization algorithms, and verifying analysis sensitivity. In order to avoid

the lack of rationality of the simulation parameter calibration results, Gong et al. [12] tried to use the mean square error method to verify the flow rate and speed, respectively. Although the calibration parameters of traffic data for HOV (high occupancy vehicles) lanes and three interacting bottlenecks were used, the driving behavior parameter was not considered in the analysis [13]. Bartin et al. [14] presented the process of building a large-scale traffic simulation model with the use of multi-source data for its calibration process via a real-world case study. With the advancement of intelligent optimization techniques such as genetic algorithms, calibration methods can carry out the systematic optimization of all parameters, avoiding the disadvantage of overgeneralization in the experimental design method. Essa and Sayed [15] indicated that microsimulation models shouldn't evaluate safety without proper model calibration. In addition, Essa emphasized VISSIM parameter can usually be transferable between two locations, because the dynamic parameters provide a better correlation between field-measured and simulated conflicts than using the default VISSIM parameters. However, more work is still needed to confirm the specific range of simulation parameters. Gallelli et al. [16] pointed out that although the simulation calibration process is very important, simulation parameters are usually performed at default values due to the difficulty of data collection and the lack of existing simulation guidelines.

The calibration process of the traffic microsimulation model has gradually stabilized after nearly 30 years of development, but analysis of simulating parameter threshold value still needs to be further studied. We generally select default parameters to finish the calibration during simulation. However, the value being calibrated shows a large range of fluctuation. In Summary, the primary purpose of the current microsimulation parameter calibration is to optimize all the parameters. When there are too many parameters, there is more than one optimal solution; then the characteristic analysis of the optimal solution is vague, and the influence between parameters cannot be judged. Therefore, this paper proposes to divide the traffic microsimulation parameters into two parts: basic parameters and model parameters. Basic parameters can be obtained by engineering measurement methods, including maximum vehicle acceleration, acceptable acceleration, maximum deceleration, acceptable deceleration, and desired speed, but not all basic parameters need to be calibrated. Model parameters refer to parameters which only act on specific driving behaviors. For example, the lane change gap indicator only acts on the lane changing behavior process and has nothing to do with other vehicle motion behaviors. This method determines the specific range of the basic parameters by engineering measurements, which breaks through the limitations of the high latitude parameter calibration in the previous microsimulation model. It has good engineering interpretation ability.

During the current calibration work, researchers usually investigate vehicles' desired speed, but they adopt default parameters on vehicles acceleration. Taking VISSIM as an example, the vehicle's default starting acceleration is 3.5 m/s^2 and the default deceleration is -2.8 m/s^2 . These two parameters are related to the vehicle's own power performance and the driver's driving experience. The default acceleration reaches the level of the racing car during the start-up phase, which has a strong push back and discomfort. However, vehicle can't achieve the vehicle testing performance

Table 1 Statistics of the vehicles average acceleration (m/s^2)

Vehicle model	Sample size	Max	Min	Mean	Standard deviation
POLO 2014	9	2.62	1.98	2.31	0.224
Santana 2015	10	3.05	2.24	2.46	0.242
Passat 2014	16	3.7	2.67	3.12	0.351
All samples	35	3.7	1.98	2.72	0.473

in reality. In the vehicle performance evaluation, we analyzed the vehicle's acceleration time from 0 to 100 km/h as the main index for analysis. In general, a 1.6 T car needs about 11–14 s to accelerate from 0 to 100 km/h (i.e., 2.52, 1.98 m/s^2). The 2.0 T medium-sized car's 100 km acceleration time is about 7–8 s (i.e., 3.47, 3.96 m/s^2). Compared with three economical vehicles that have the selling price between 100,000 and 200,000 Chinese Yuan (i.e., 14,677.50–29,355.00 US Dollar), like VOLKSWAGEN GOLF, SANTANA, and PASSAT, the experiment calculated their average maximum acceleration [17]. The vehicle detail indexes are shown in Table 1. Simply put, default parameters differ quite markedly from actual vehicle performance.

Therefore, acceleration and deceleration can be divided into adaptable and acceptable acceleration/deceleration, unadaptable but acceptable acceleration/deceleration, unadaptable and unacceptable acceleration/deceleration, and maximum acceleration/deceleration. Maximum acceleration and deceleration are related to the vehicle's dynamic performance; the other three types are related to the driver's own actions. The three types of acceleration/deceleration correspond to the different perceptions of the passengers in the car. If the passenger does not have the feeling of leaning forward or backward during the experiment, the state is adaptable and acceptable, corresponding to the adaptable and acceptable acceleration/deceleration. If the passenger has the feeling of leaning forward and reclining but does not need to complete the balance by hand and leg strength, the state is unadaptable but acceptable, corresponding to the unadaptable but acceptable acceleration/deceleration. If the passenger has a feeling of leaning forward or backward, it is difficult for the individual to keep one's own balance without the help of the hands or feet; the state is unadaptable and unacceptable, corresponding to unadaptable and unacceptable acceleration/deceleration.

The desired speed is usually the highest safe speed for a driver during the driving process. This study selects the 85th percentile speed of the vehicle as the desired speed. Before the vehicle reaches the desired speed, an adaptable and acceptable acceleration or unadaptable and acceptable acceleration is chosen for operating the vehicle. After the vehicle reaches the desired speed, the driver selects an adaptable and acceptable acceleration to maintain the vehicle state. In order to ensure driving safety, the driver chooses the maximum deceleration and deceleration to operate the vehicle, and the vehicle maximum deceleration can reach 10 m/s^2 in emergency.

2 Experimental Design

Simulation parameters are closely related to the simulated vehicle speed. At present, the calibrations of the microsimulation model focus on the optimization of the simulation parameters, but it lacks measured vehicle road data. As a result, the calibration result of the simulation model does not have a practical meaning, and it also reduces the portability of calibration results.

The experiments were tested with Video VBOX Pro products, which can update all GPS parameters. These products can also record the vehicle environment and obtain vehicle motion data simultaneously in real time.

The experiments aimed to obtain driving behavior characteristics. The main investigations included the maximum acceleration, maximum deceleration, acceptable acceleration, acceptable deceleration and the relation between speed and acceleration. Two experiments in this study intended to obtain the basic parameters in the simulation, which were carried out with the Volkswagen Bora 1.6 L automatic transmission vehicle equipped with the Video VBOX Pro. This research chose 20 participants with skilled driving experience in two experiments. Participants drive at least 10,000 km per year and their ages between 25 and 40.

Experiment 1

This experiment aimed to obtain the distribution interval of adaptable and acceptable acceleration/deceleration, unadaptable but acceptable acceleration/deceleration, unadaptable and unacceptable acceleration/deceleration.

This research chose a closed road 500 m in length at Beijing University of Technology. The 20 participants each accelerated to 60 km/h and then decreased speed to the stop state with different acceleration and deceleration on the road. During the process, two experimental testers recorded the feeling of each acceleration and deceleration state as passengers in the vehicle, which was classified as adaptable and acceptable, unadaptable but acceptable, or unadaptable and unacceptable. Each participant performed 5 tests, respectively.

Experiment 2

This experiment aimed to examine the relationship between vehicle speed and acceleration. This experiment was tested on the urban road, including Nanmofang Rd., Dawang West Rd., Jianguo Rd. and Third Ring Rd. in Beijing. The designed route includes three road grades: expressway, major arterial road, and minor arterial road, which help the study of desired speed on different roads. Table 2 shows the specific geometric parameters. Participants drove the vehicle on the selected path during off-peak hours. Each participant drove one cycle path using the Video VBOX Pro device to record vehicle data during the experiment.

Table 2 Geometric characteristics of the segments used in the road test

Rd name	Road grade	Number of lanes	Speed limit (km/h)	Length (km)
Third Ring Rd	Expressway	4	80	2.7
Nanmofang Rd	Secondary trunk road	3	50	1.3
Dawang West Rd	Arterial road	3	60	2.7
Jianguo Rd	Expressway relief road	3	60	1.3

3 Results and Analysis

An overwhelming amount of noisy data exists when using GPS to obtain vehicle operating data. Therefore, raw data need to be optimized using Performance Tools, which is the Video VBOX Pro device companion software. In addition, Performance Tools can capture the acceleration/deceleration interval of the data according to the user's needs, which is very convenient for obtaining the vehicle's acceleration and deceleration data.

Experiment 1 Data Processing Method

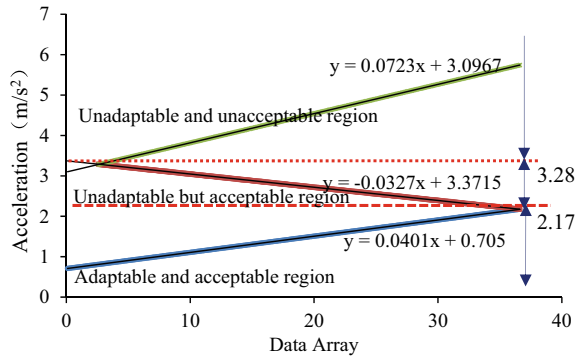
The subjective evaluation method was used in Experiment 1. Every experiment was investigated by two testers regarding the feeling of acceleration and deceleration of vehicle movements, adaptable and acceptable acceleration/deceleration, unadaptable but acceptable acceleration/deceleration, or unadaptable and unacceptable acceleration/deceleration. In the process of the vehicle starting from a standstill and entering the parking state, the vehicle's inertia could cause a sudden change in the experimental results. Therefore, Performance Tools set the acceleration interval (5–30 km/h) and the deceleration interval (30–5 km/h) to obtain the vehicle's maximum acceleration/deceleration during the experiment.

The experiment testers generally had the same perceptions, but the evaluation in the different acceleration/deceleration states differed in each of the acceleration and deceleration state according to the characteristics of the vehicle tester. This study used a no-signal intersection method by which the vehicle passed through the critical gap to determine the extent of the area in different states.

The acceleration value in the adaptable and acceptable state and the unadaptable and unacceptable state were counted in ascending order, the acceleration value in the unadaptable but acceptable state was counted in descending order, and then the series was fitted with a linear relationship. The fitting curve divided the entire acceleration interval into three parts, as shown in Fig. 1, which are $0\text{--}2.17\text{ m/s}^2$, $2.17\text{--}3.28\text{ m/s}^2$, and $> 3.28\text{ m/s}^2$.

Similarly, the deceleration value in the adaptable and acceptable state and the unadaptable and unacceptable state was counted in ascending order, the deceleration value in the unadaptable but acceptable state was counted in descending order, and then the series was fitted with a linear relationship. The fitting curve divided the

Fig. 1 Fitting figure of acceptable acceleration threshold judgement



entire deceleration interval into three parts, as shown in Fig. 2, which are 0 to -2.25 m/s^2 , -2.25 to -3.48 m/s^2 , and $< -3.48 \text{ m/s}^2$.

In conclusion, the distribution interval of adaptable and acceptable acceleration/deceleration, unadaptable but acceptable acceleration/deceleration, and unadaptable and unacceptable acceleration/deceleration is shown in Table 3.

Experiment 2 Data Processing Method

Fig. 2 Fitting figure of acceptable deceleration threshold judgement

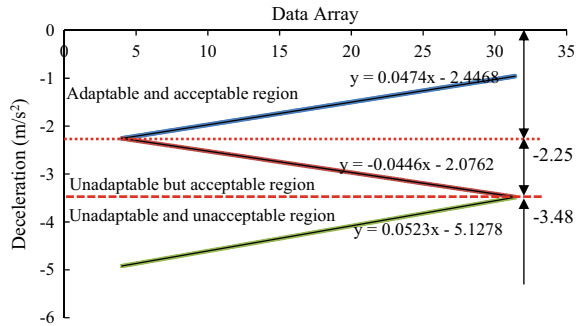


Table 3 Speed fitness interval distribution of acceleration and deceleration

Acceleration value (m/s^2)	Adaptable and acceptable	Unadaptable but acceptable	Unadaptable and unacceptable
Distribution interval	0–2.17	2.17–3.28	> 3.28
Deceleration value (m/s^2)	Adaptable and acceptable	Unadaptable but acceptable	Unadaptable and unacceptable
Distribution interval	0 to -2.25	-2.25 to -3.48	< -3.48

Table 4 Measured speed statistics

Speed statistics (km/h)	50%	85%	90%	Max
Nanmofang Rd	41.9	50.5	51.8	55
Xidawang Rd	45.3	55.7	57.2	62.9
Third Ring Rd	64.3	71.1	72.3	77.4

Due to the influences caused by equipment and other random factors, this research collected 17 sets of valid driving data from a total of 20 driving experiments. One of the three samples occurred a car accident during the test, it led the significant defects in the data. While the other two samples were encountered the severe traffic congestion, it is difficult to show a complete operation process.

Experiment 2 Data Processing Method

Determine the desired speed

When the vehicle speed is higher than 30 km/h, the traffic flow can be considered to be in a non-congested (i.e. car-following) state. The vehicle was difficult to operate in the state of free flow during the actual investigation; therefore, this research chose the data with vehicles' speed above 30 km/h for statistical analysis. Table 4 shows the cumulative speed statistics on the investigated road. Although the desired speed was greatly affected by the driver's own characteristics, the 85th percentile speed can be used as the desired speed. In this research, the 85th percentile average speed of drivers was taken as the desired speed, which was used as the range of vehicle acceleration and deceleration.

Acceleration and deceleration analysis

When the vehicle does not reach the desired operating state, the driver will accelerate to the desired speed. Two types of vehicle acceleration operation can occur during the process: free starting acceleration and starting acceleration with a following vehicle. Free starting acceleration is the focus of this research; another is the result of vehicle interactions. The study sets the acceleration data interval for Nanmofang Rd., Dawang West Rd. and the Third Ring Rd. to 0–50.5 km/h, 0–55.7 km/h, 0–71.1 km/h, and combines video information to obtain continuous acceleration data in continuous vehicles. As a minor urban arterial road, Nanmofang Rd. has many signalized intersections, which makes it difficult to obtain the free acceleration process. The vehicle has already entered the high-speed motion state when it enters the Third Ring Rd. Therefore, this research focuses on the free starting process of vehicles on Dawang West Rd. Figure 3 shows the scatter plot of the starting speed and acceleration relationship on Dawang West Rd.

In the past, data classification was usually carried out by cluster analysis. However, the acceleration is affected by many factors in the process of vehicle acceleration, and the acceleration changes at different speeds are very different. The classification results do not have a specific meaning if we only cluster raw data. According to the driving experience, the vehicle acceleration process can be divided into three

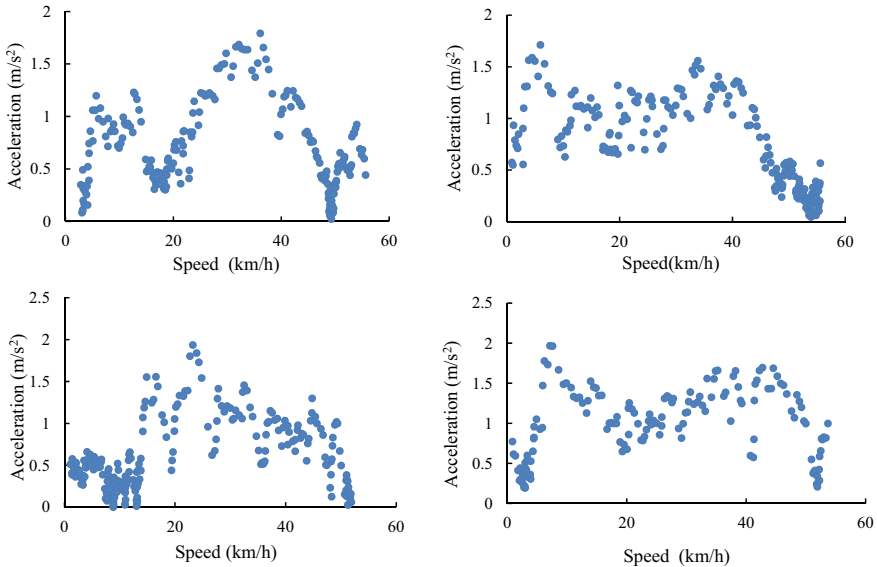


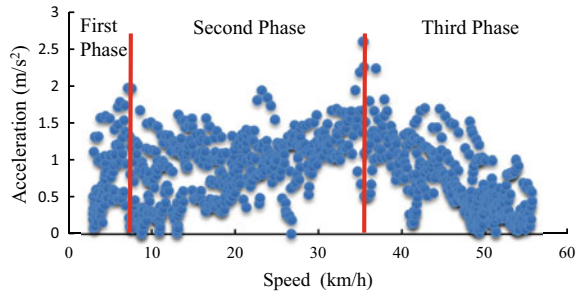
Fig. 3 Sample of relationship between vehicle speed and acceleration

phases: starting in gear and lightly stepping on the accelerator pedal, increasing pressure on the accelerator pedal to accelerate speed, and reducing pressure on the accelerator pedal to maintain the speed appropriately. Similarly, the acceleration process can be divided into three phases for characteristics. The first is the starting phase, in which the vehicle increases speed with slight acceleration. After that, vehicle operates in accordance with the acceptable acceleration at increasing phase, the speed is rapidly increased. When the vehicle speed approaches the desired speed, acceleration gradually decreases, meaning the vehicle enters the stabilizing phase, and the vehicle speed tends to be stable. Therefore, it is necessary to determine the relationship between the threshold speed at each phase and the relationship between the acceleration and speed, and then determine the relationship between the speed of the overall acceleration process and the acceleration.

This research divides the data into three phases with the peak as the threshold. As can be seen in Fig. 4, there is no exact mathematical relationship between speed and acceleration. Under certain speed conditions, the acceleration can be randomly selected within a certain area. In Geometry, the envelope of family of curves is a curve in which each point touches tangentially one of the curves of the family. The acceleration value interval was described, adopting the form of an envelope.

According to statistics, the critical speed between Phase 1 and Phase 2 is 7.2 km/h, and the critical speed between Phase 2 and Phase 3 is 35.4 km/h. It determines the relationship between the acceleration and speed in the starting phase, increasing phase and stabilizing phase. The measured data contains large numbers of random factors. Therefore, the raw data cannot be used directly. In order to determine the relationship

Fig. 4 Sum of relationship between vehicle speed and acceleration



between speed and acceleration precisely, this research used the mean value method of the acceleration under the same speed condition to reduce the influence of random factors. The data processing results are shown in Fig. 5.

The linear regression method is applied to fit the processed data for the first phase, and the data distribution envelope is solved according to the sample edge points, as shown in Fig. 6. It can obtain the first stage fitting equations: $y_1 = 0.1832x + b_1$, $b_1 \in (-0.3989, 0.1811)$.

The second and third phases are analyzed and fitted respectively according to the same method. The second and third phases of the fitted equations can be obtained separately: $y_2 = 0.0256x + b_2$, $b_2 \in (-0.0803, 0.9931)$, $y_3 = -0.0644x + b_3$, $b_3 \in (3.1172, 4.0802)$, The analysis results are shown in Figs. 7 and 8 respectively.

Fig. 5 Sum of relationship between vehicle speed and acceleration after average processing

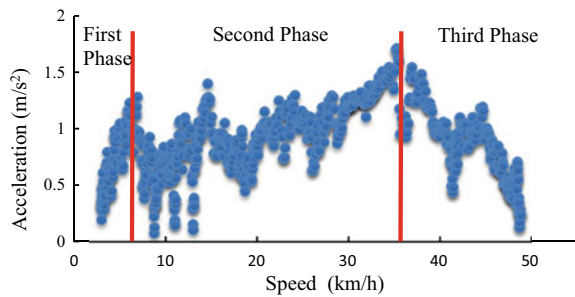


Fig. 6 Data regression analysis of the first phase

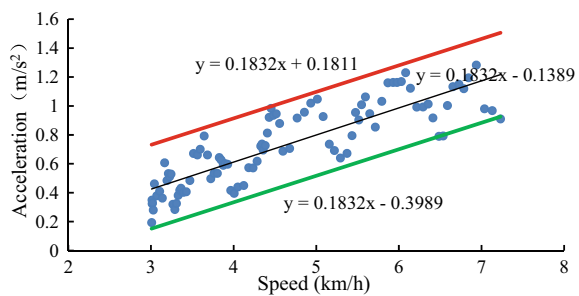


Fig. 7 Data regression analysis of the second phase

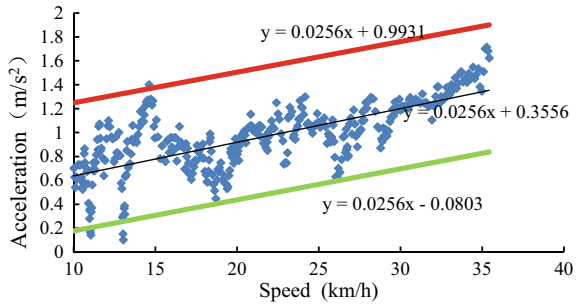
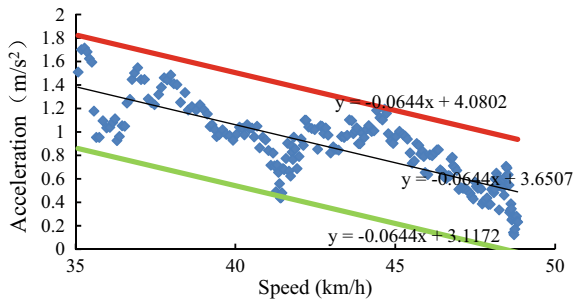


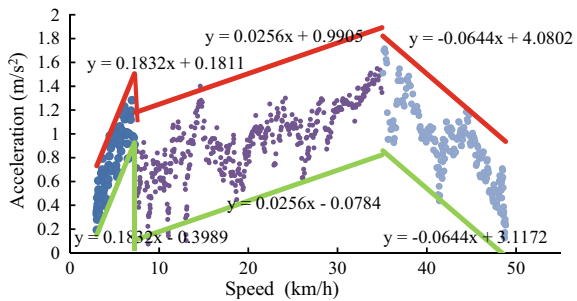
Fig. 8 Data regression analysis of the third phase



This research chose the linear regression method to fit processed data in each phase using the same fitting slope to solve the data distribution envelope according to the sample edge points. Finally, the relationship between speed and acceleration during the vehicle acceleration is shown in Fig. 9. At a given speed, acceleration value can be selected randomly in the envelope function.

The same data performance characteristics apply to expressways and minor arterial roads, but a critical threshold division on the boundary is different. The higher the desired speed, the greater critical threshold for the second and third phases. Due to the small sample size, it is not sufficient to accurately describe the relationship between speed and acceleration on the expressway, but it can be determined that the second and third phase speed critical threshold is between 35.2 and 46.0 km/h.

Fig. 9 Regression relationship between vehicle speed and acceleration



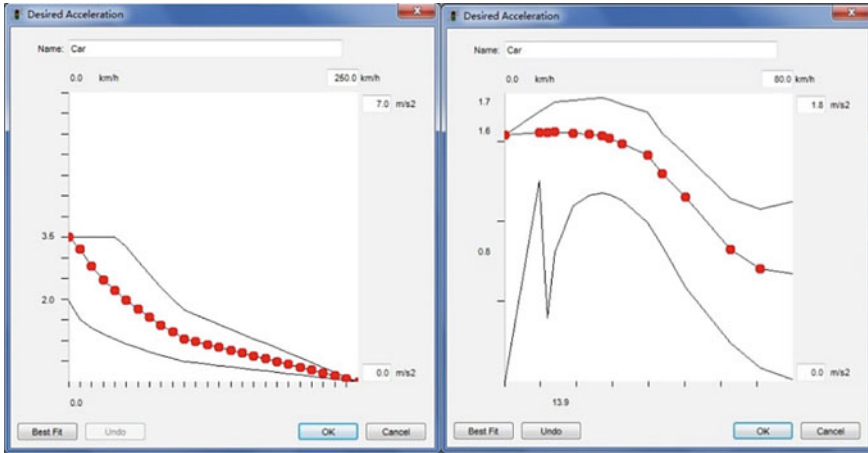


Fig. 10 Default acceptable acceleration versus adjusted value

In summary, three phases have its own characteristics. In starting phase ($V < 7.2$ km/h), vehicle is needed large power to run. The speed is low and the acceleration is large, the acceleration-speed slope is $k = 0.1832$; In increasing phase ($V = 7.2\text{--}35.4$ km/h), in order to ensure the environment is in a safe state, the acceleration is reduced a little and then gradually increased, the overall slope of the acceleration-speed in this phase is $k = 0.0256$. In stabilizing phase ($V > 35.4$ km/h), the speed is relatively stable, and the acceleration-speed slope is approximately equal to zero. It determines the fine relationship between acceleration and speed in base parameter and optimizes the relationship between default acceleration and speed in the simulation software.

As can be seen in Fig. 10, the default acceleration curve in VISSIM differs greatly from the acceleration curve obtained by actual investigation. If default acceleration is used for simulation, the acceleration will be too great and the delay value will be too small, resulting in distortion. Therefore, this research could reproduce traffic scenes more realistically.

4 Conclusions and Discussions

This research measured the basic parameters and analyzed the maximum acceleration and deceleration according to vehicle movement characteristics. However, in order to ensure driving comfort, drivers often choose acceptable acceleration/deceleration during the driving. The study performed a qualitative analysis of acceptable acceleration/deceleration in Experiment 1 and developed an obtaining engineering experiment method of acceptable acceleration/deceleration in the traffic microsimulation model. In addition, the choosing and changing of acceleration is the foundation

of traffic microsimulation. Based on the analysis in Experiment 2, this research proposed the envelope area that centered on vehicle speed and acceleration fit line as the speed and acceleration model, and then it further provided a specific function line. Although the measurement calibration process is time-consuming and laborious, this method can give a more accurate parameter threshold range, which also makes the microsimulation closer to reality.

This research can provide reference for updating parameters of the simulation software and calibration parameters for users in the future. It was tested on the urban road, therefore, the threshold values of urban roads can be calibrated using the results. However, only one of the acceleration and deceleration modes for a validation illustration was tested, and the sample size was also limited. To improve the accuracy and effectiveness of calibration parameters, the exploration of acceleration and deceleration modes with a larger sample size is expected. This research can help researchers to get more precise parameter threshold, specifically for acceleration-speed parameters thresholds under urban road. In the future, the experiment will focus on highways and rural roads.

References

1. Akbar M, Khan R, Khan MT, Alam B, Elahi M, Wali B, Shah AA (2018) Methodology for simulating heterogeneous traffic flow at intercity roads in developing countries: a case study of university road in Peshawar. *Arab J Sci Eng* 43(4):2021–2036
2. Benekohal R (1991) Procedure for validation of microscopic traffic flow simulation models. *Transp Res Rec J Transp Res Board* 1320:190–202
3. Hellinga BR (1998) Requirements for the validation and calibration of traffic simulation models. *Proc Can Soc Civ Eng* 4:211–222
4. Brockfeld E, Kühne RD, Wagner P (2004) Calibration and validation of microscopic traffic flow models. *Transp Res Rec J Transp Res Board* 1876:62–70
5. Park BB, Schneeberger JD (2003) Microscopic simulation model calibration and validation: a case study of VISSIM for a coordinated actuated signal system. In: *Proceedings of 82nd transportation research board annual meeting*, Washington, D.C., Jan 2003
6. Azam M, Puan OC, Hassan SA, Mashros N (2019) Calibration of microsimulation model for tight urban diamond interchange under heterogeneous traffic. *IOP Conf Ser Mater Sci Eng* 012077
7. Guo Y, Essa M, Sayed T, Haque MM, Washington S (2019) A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia. *Transp Res Part C Emerg Technol* 101:96–110
8. Gomes G, May A, Horowitz R (2004) A microsimulation model of a congested freeway using VISSIM. In: *Proceedings of the 83rd transportation research board annual meeting*, Washington, D.C., Jan 2004
9. Merritt E (2004) Calibration and validation of CORSIM for Swedish road traffic condition. In: *Proceedings of the 83rd transportation research board annual meeting*, Washington, D.C., Jan 2004
10. Rong J, Ma J, Ren F, Liu X (2001) Solving capacity using traffic imitation model. *J Beijing Univ Technol* 27(1):4–7
11. Ciuffo B, Punzo V, Torrieri V (2007) A framework for the calibration of microscopic traffic flow models. In: *Proceedings of 86th transportation research board annual meeting*, Washington, D.C., Jan 2007

12. Gong L et al (2016) Calibrating microscopic traffic simulation models using speed-flow data. In: *Advances in transportation studies*, vol 39
13. Gabriel G (2014) A microsimulation model of congested freeway using VISSIM. In: *Proceedings of the 93rd annual meeting of transportation research board*, Washington, D.C., Jan 2014
14. Bartin B, Ozbay K, Gao J, Kurkcu A (2018) Calibration and validation of large-scale traffic simulation networks: a case study. *Procedia Comput Sci* 130:844–849
15. Essa M, Sayed T (2017) Transferability of calibrated microsimulation model parameters for safety assessment using simulated conflicts. *Accid Anal Prev* 84:41–53
16. Gallelli V et al (2017) Investigating the transferability of calibrated microsimulation parameters for operational performance analysis in roundabouts. *J Adv Transp*
17. Chang X, Rong J, Zhou C et al (2016) Research on vehicle running characteristics based on driver's perception. *J Transp Inf Saf* 01:51–56

A Comparative Study on Institutions and Mechanisms for Investigating and Handling Maritime Accidents in China and Major Developed Countries



Minglu Ma and Lili Jiang

Abstract Maritime accident investigation is the most effective tool used for controlling maritime risks, and is also a managing procedure indispensable to preventing against and reducing maritime accidents. China could learn much from developed countries in Europe and America, which boast abundant experience in establishing institutions and mechanisms for maritime accident investigation, in order to advance establishment of institutions and mechanisms for maritime accident investigation in China. Specifically speaking, with a comprehensive comparative research being made on respective advantages and disadvantages of developed countries in this regard, China could absorb advanced management ideas and methods from those countries, improve the institutions for investigating and handling maritime accidents, further enhance the capacity to investigate and handle maritime accidents, safeguard maritime traffic safety and reduce maritime accidents.

Keywords Maritime accident investigation · Maritime accident handling · Maritime accident mechanism

1 Status Quo of Maritime Accident Investigation in Major Developed Countries

1.1 America

1.1.1 Management Institution for Investigating Maritime Accidents

America's maritime accident investigation is made mainly according to the Maritime Safety Manual (United States Coast Guard).

M. Ma (✉) · L. Jiang
China Waterborne Transport Research Institute, Beijing 100088, China
e-mail: maminglu@wti.ac.cn

America's maritime accident investigation is mainly conducted by the United States Coast Guard (USCG) and the National Transportation Safety Board (NTSB), which perform respective functions in this regard, with NTSB in charge of investigation of major maritime accidents. Independent from each other, these two agencies conduct investigations in accordance with their respective regulations and submit investigation reports separately.

1.1.2 Team Construction

USCG participates in maritime accident investigation through the Marine Accident Investigation Branch, one of five branches with the Maritime Safety Office subordinate to USCG. In the meantime, USCG in each region boasts a maritime safety branch where a maritime inspector works with a maritime accident investigation team.

In charge of investigating various traffic accidents in America, NTSB owns a special team called "GO TEAM", which is in essence a quick-response team, consisting of experts in various fields and being directly led by the USCG head office.

1.1.3 Procedures for Maritime Accident Investigation

(1) Investigation scope

Investigation into technical causes of maritime accidents and criminal liability is mainly conducted by officials with credentials, the qualified and other individuals, with relevant evidences being submitted to procuratorial organs. With USCG conducting a preliminary investigation, the commanding officer should inform NTSB of judgment on major maritime accidents.

(2) Investigation mechanisms

NTSB and USCG have common jurisdiction over investigation into "major maritime accidents", in which they discuss whether NTSB needs to take a part.

(3) Handling procedures

USCG conducts a preliminary investigation and informs NTSB of the results. NTSB will send one or more officers to participate in every investigation stage. Nevertheless, NTSB can decide whether to take part or not at its own discretion.

1.1.4 Investigation Reports

USCG and NTSB release respective maritime accident investigation records to the public. These two agencies, which perform respective maritime functions, conduct an

investigation into maritime accidents according to independent investigation regulations, and submit in the standard format their respective reports which include suggestions to improve maritime safety. Media could have access to copies of the investigation reports.

1.2 Britain

1.2.1 Management Institution for Maritime Accident Investigation

Britain's maritime accident investigation consists of independent safety investigation and administrative investigation, which are in the charge of the Marine Accident Investigation Branch (MAIB) and Maritime and the Coastguard Agency (MCA). Both are subordinates to the Department of Transport.

Being directly under the Department of Transport and independent from MCA, MAIB is mainly responsible for identifying causes of maritime accidents and avoiding recurrence of similar accidents. MCA is an administrative enforcement agency under the Department of Transport. It has functions similar to those of the Maritime Safety Administration of the People's Republic of China—protecting maritime safety and preventing maritime pollution. MCA's maritime accident investigation function authorized by the law merely lies in determining illegal behaviors of vessels and deciding treatments according to severity and consequences of illegal behaviors.

1.2.2 Team Construction

MAIB investigators are divided into four investigation groups, one of which is on duty each week with another group always ready for any emergency. Once a maritime accident requires a MAIB investigation, the investigation team on duty will rush to the accident site for an investigation.

All maritime accident investigators are required to attend classes for basic accident investigation skills training, and will be granted with corresponding certificates. MAIB pays much attention to health and labor protection of personnel, which is why investigators need to receive relevant health and security training before participating in an accident investigation. MAIB also provides each investigator with special equipment for labor protection and special training. Apart from that, MAIB also provides various kinds of special training in accident investigation on an annual basis, usually for all staff, who could share special knowledge and experience.

1.2.3 Procedures for Maritime Accident Investigation

(1) Investigation scope

According to the Merchant Shipping Act released in 2005 in Britain, a maritime accident investigation is aimed at identifying the objective environment and causes of the accident, while determining legal liability or holding any party accountable is not the goal of a maritime accident investigation.

Maritime accidents, which MAIB is in charge of investigating, include accidents that happen in any vessel within Britain's territorial waters and accidents happening to any British vessel in waters around the world. Furthermore, maritime accidents that MAIB needs to investigate also incorporate both maritime traffic accidents, and any safety accidents and perils at sea.

MAIB classifies maritime accidents into the following three kinds: maritime casualty, vessel accidents and maritime incidents (accidents are likely to happen, but do not happen). Based on analysis of maritime accidents, different levels of investigations are conducted: no investigation, managing investigation, preliminary investigation and full investigation. Different methods are adopted for different levels of investigations, reaching different results.

(2) Investigation mechanisms

Britain's maritime accident investigation constitutes safety investigation and administrative investigation, which are in the charge of MAIB and MCA. Britain's investigation institutions avoid combination of punishment and objective technical investigation from the perspective of safety, and boast the relatively complete investigation procedures, the complete system for report release and feedback, proper management of specialized personnel, diversified and highly-efficient teams and so on. However, division of labor in the British institutions might lead to repeated investigations, waste of resources and increased cost of human and material resources.

(3) Handling procedures

MAIB

1. Maritime accident reporting

After a maritime accident is reported, the on-duty coordination officer (DCO) will make a preliminary assessment and log onto the accident report memorandum (ARM).

2. Preliminary assessment

After receiving the ARM, the DCO should make a timely assessment of the nature and severity of an accident, deciding whether an investigation should be made.

3. Investigation

Once a safety investigation is determined, the DCO will inform parties concerned of the MAIB's decision to launch a safety investigation. When

deciding not to launch a safety investigation, the DCO must clarify necessary measures apart from the minimum information required by MAIB.

4. Deployment

A no-entry notice should be issued in order to collect more evidences during a field investigation, and interested parties should be notified.

5. Monitoring and assessment

Creating a uniform understanding of the accident scenario; improving accident analysis; marking investigation routes; identifying tests or experiments that will be conducted; posting notice; determining time for assessment.

6. Advice on safety management

Advice on safety management must be necessary, effective, pragmatic and purposeful; identifying safety issues that need to be resolved could lay a foundation for countermeasures.

MCA

MCA's maritime accident investigation follows no fixed modes, and its investigation result will not be accessible by the public.

1.2.4 Investigation Report

Accident investigation advice is the main fruit from MAIB. MAIB mainly notifies of accident advice the parties concerned including the Department of Transport, MCA, the International Chamber of Shipping, and other governmental agencies or NGOs. MAIB could also urge and advise on compulsory enforcement of law.

(1) Investigation levels

Based on different maritime accident investigators, investigation could be divided into the following three levels: investigation by a chief investigator; investigation by an investigator; administrative investigation.

(2) Report types

MAIB' s investigation reports could be classified into three kinds:

1. Standard report. This type of report applies in the following situations: when an investigation might bring about the first and second kinds of management advice, and involves relevant policies and rules pertaining to the government or industrial institutions; or when an investigation might attract attention from the public or industries, or involve important safety issues.
2. Concise report.
3. Simple report. This type of report will only available on the website of MAIB.

(3) Advice types

Accident investigation advice in Britain could be classified into three types as follows: important advice on formulating new laws or modifying existing policies; advice on governmental agencies or organizations altering existing management regulations or forcing implementation of optimal management regulations; advice for individual shipowners or enterprises.

1.3 Japan

1.3.1 Management Institution for Maritime Accident Investigation

Investigation into maritime traffic accidents in Japan is mainly the responsibility of Japan Transport Safety Board (JTSB) and the Japan maritime Accident Tribunal (JMAT). In charge of investigating causes of a maritime traffic accident, JTSB also puts forward preventive measures in order to prevent recurrence of similar accidents, while JMAT punishes accountable crew in an accident.

JMAT mainly imposes administrative penalty on administrative bodies, without criminal justice or civil penalty being involved.

1.3.2 Mechanisms for Maritime Accident Investigation

(1) Investigation scope

Investigation into causes of maritime traffic accidents and determination of accident liability are within the line of duty of different governmental departments in Japan. Specifically speaking, investigation into causes of maritime accidents is completed by an investigation group headed by a maritime accident investigator and is aimed at identifying the technical causes for maritime accidents; liability identification is the responsibility of a group headed by a judge.

(2) Investigation mechanisms

Maritime accident investigation and approval are guided by the principle of separation of powers and check, and adopt the debate system and the right relief system.

(3) Handling procedures

JTSB's procedures for investigation into maritime accidents are listed as follows:

1. Maritime accident reporting. After a preliminary judgment, JTSB will send a corresponding accident investigator to investigate causes of a maritime accident.
2. Investigation into causes of the maritime accident. When a major maritime traffic accident or incident happens, JTSB will send an investigator in the head

office responsible for vessel accidents to investigate causes of the accident or incident. Maritime accidents or incidents other than major maritime accidents are in the charge of a local maritime accident investigator.

3. Compilation, review and approval of maritime accident reports.
Maritime accident investigation reports compiled by accident investigators are submitted to corresponding governmental agencies for review and approval.
4. Investigation results being released to the public.

1.4 South Korea

1.4.1 Management Institution for Maritime Accident Investigation

Subordinate to the Ministry of Oceans and Fisheries, the Korea Maritime Safety Tribunal (KMST) boasts four local maritime safety tribunals, which perform functions in maritime accident investigation within respective jurisdictions, and have the right to investigate various vessel accidents except for unilateral accidents involving military vessels and public service ships. KMST adopts the same institution as that of the Seoul Administrative Court. The review and approval procedures in the institution are similar to those in a criminal or civil court, and are independent.

As with Japan, South Korea also demarcates global waters.

1.4.2 Team Construction

In South Korea, judges and chief investigators report to corresponding chief justices, and a chief investigator takes charge of a corresponding investigation division and administrative department. Local maritime safety tribunals have the responsibility to investigate and judge maritime traffic accidents (including fishery vessel accidents) within respective jurisdictions.

1.4.3 Procedures for Maritime Accident Investigation

(1) Investigation scope

KMST makes a maritime accident investigation mainly in order to identify main causes of maritime accidents, determine mistakes and improper behaviors of relevant crew and responsible personnel, and provide safety advice, thus preventing similar accidents from happening again.

(2) Investigation mechanisms

KMST adopts the system employed by the Seoul Administrative Court. The system boasts review and approval procedures utilized by a criminal or civil court, and is

independently implemented. Each of local maritime safety tribunals consists of a chairman, judge, investigator and administrative staff, who work closely with each other to present clear procedures.

(3) Handling procedures

When a maritime accident investigator gets to know or reports from shipping agencies, police or local officials reveal that a maritime accident has happened, an investigator will immediately summon accident-related people and inquire about an exact site of vessels and the accident. An investigator could also require parties concerned to submit documents, which mainly include accident reports and original records, and ask the police station to provide reports and data.

1.4.4 Investigation Reports

After the Korea Central Shipwreck Tribunal completes an accident investigation with an investigation report being also compiled, the chief investigator should release the report in an official newspaper in the form of notice, which should also be submitted to Minister of the Ministry of Oceans and Fisheries. If necessary, the report should also be published in a newspaper. Much attention should be paid to advice in the investigation report, based on which corresponding actions should be taken.

2 Status Quo of Maritime Accident Investigation and Handling in China

2.1 The Investigation Institution

In China, maritime accident investigation bodies and division of supervision labor are complex. At present, investigation into “maritime accidents” are made by various agencies or organizations which have different investigation rights. The Maritime Traffic Safety Law of the People’s Republic of China is the major legal basis for maritime accident investigation at present, and is also the only law about maritime traffic safety since establishment of the People’s Republic of China. The Regulations of the People’s Republic of China on the Investigation and Handling of Maritime Traffic Accidents is a document about specific actions. China’s maritime safety administrations are the main forces for investigating and handling maritime traffic accidents, and bureaus directly under the maritime safety administrations own maritime safety offices, which establish mechanisms for joint investigation into vessel pollution accidents and maritime traffic accidents, laying a foundation for further regulating maritime accident investigation and handling.

Maritime traffic accident investigation teams consist of investigators who are transferred from departmental maritime safety administrations, maritime safety

administrations directly under departmental maritime safety administrations, and provincial maritime safety administrations. Departmental maritime safety administrations could directly transfer investigators from maritime safety administrations directly under departmental maritime safety administrations and provincial maritime safety administrations, to an investigation team; maritime safety administrations directly under departmental maritime safety administrations, or provincial maritime safety administrations could send investigators within jurisdictions to form an investigation group.

2.2 Team Construction

Construction and management of maritime accident investigator teams are of much importance to ensuring efficient investigation and handling of maritime accidents. The Regulations on Management of Maritime Accident Investigators clearly stipulates maritime accident investigators' levels, classifications, training, exams, certification, evaluation and so on, and acts as the main basis for selecting maritime accident investigators in China. China's maritime accident investigators are classified into senior investigators, intermediate investigators, and assistant investigators of maritime accidents. Senior and intermediate maritime accident investigators are divided into two kinds: foreign investigators and domestic investigators. All maritime accident investigators need to receive different levels of training and pass various exams in a bid to obtain official certificates, and maritime accident investigating personnel without a certificate are not allowed to be in charge of maritime accident investigation and handling.

2.3 Procedures for Investigating and Handling Maritime Accidents

Summary procedure: minor accidents with an evident scenario; general procedures: general procedures include reporting an accident, accepting a case, investigating an accident, closing a case, maritime mediation (if any), filing a case and other procedures.

Departmental maritime safety administrations establish mechanisms for transferring maritime accident investigators and expert counseling, which mainly include maritime accident investigation expert committee, foreign maritime accident investigation task force, and sub-committee on maritime traffic accident investigation and handling. Departmental maritime safety administrations could transfer investigators from any maritime safety administrations, forming an investigation team; maritime safety administrations directly under departmental maritime safety administrations, or provincial maritime safety administrations could send any investigator

within jurisdictions to form an investigation team. Maritime safety administrations at various levels pay much attention to establishment of maritime accident institutions, thus further improving institutions for case accepting, investigation, analysis, case closing, investigator management and so on, and ensuring that maritime safety administrations could perform respective functions effectively.

3 Comparative Analysis

(1) Agencies undertaking maritime accident investigation

At present, in most developed countries, maritime traffic accident investigation is usually made by accident investigation agencies relatively independent from the department of transport, such as USCG in America and MAIB in Britain. That provides a basic guarantee for authenticity and impartiality of accident investigation. In contrast, safety investigation and administrative investigation in China are combined, and are in the charge of maritime safety offices or a department under maritime safety administrations, with independent agencies for maritime accident investigation and handling being established. At present, China's investigation institution indeed makes for comprehensive utilization of technologies and reduced enforcement costs, but has adverse influence on identifying root causes of accidents and development of maritime accident investigation.

(2) Purpose of accident investigation

In countries such as America, Britain and Japan, cause investigation and liability investigation for maritime traffic accidents are performed by different governmental departments: investigation in causes of maritime accidents is made by an investigation team headed by a maritime accident investigator, and is aimed at identifying causes of accidents from the technical perspective; accident liability investigation is made by a team headed by a judge. In China, however, investigation into accident causes and accident liability investigation are completed by the same accident investigation team: investigation into accident causes is made along with analysis of liabilities for the accident. The fact that accident investigation conclusions could hardly be applied in reality leads to a shift of the main goal of accident investigation from preventing recurrence of similar accidents to determining responsibilities borne by both parties in an accident. Furthermore, because investigation teams are formed within a very short time, some accident investigators could be easily influenced by organizational or individual emotional elements when determining accident liability, thus damaging impartiality.

(3) Application of accident investigation results

Maritime traffic investigation agencies in developed countries all establish a quite complete track and feedback system: tracking and monitoring every safety advice until it is adopted and implemented, and achieves the expected effect. For instance, a

period of time after putting forward safety advice, USCG in America inspects relevant departments, or requires relevant enterprises to report advice implementation to it in writing within a certain period of time; Britain requires the following: within four weeks after receiving a report, relevant departments pertaining to safety management advice should submit an overhaul plan to MAIB based on advice, with a timetable being attached.

However, in China, maritime accident investigation reports are not open and transparent enough: only major maritime traffic accident investigation reports are required to be made open to the public, while reports of investigation into maritime accidents below major maritime accidents are not forced to be open. Judging from some safety advice included in China's maritime traffic accident investigation reports, it is clear that most of safety advice is not purposeful enough, and as result, enterprises fail to follow those advice in the process of overhaul.

4 Advice

- (1) Improving laws and regulations. Necessary changes should be made to the Maritime Traffic Safety Law of the People's Republic of China and the Regulations of the People's Republic of China on the Investigation and Handling of Maritime Traffic Accidents. For instance, maritime safety investigation should be added to the Regulations of the People's Republic of China on the Investigation and Handling of Maritime Traffic Accidents: situations where maritime safety investigation is necessary should be summarized in order to create formal rules.
- (2) Setting specialized agencies for maritime safety investigation. Specialized offices for maritime safety investigation could be added to existing maritime safety agencies, and consist of existing maritime accident investigators, in order to ensure quality of maritime safety investigation.
- (3) Establishing new-type mechanisms for maritime accident investigation. The relations between independent maritime accident investigation agencies and maritime accident investigation teams should be clarified, duties of maritime accident investigation agencies and (temporary) maritime accident investigation teams should be differentiated, and the relations between technical investigation reports and administrative penalty should be clarified.
- (4) Establishing a system for release of publically-available investigation reports, supervision and overhaul.

Acknowledgements I am grateful to my colleagues Xu Liansheng and Chen Fengyun, who helped me with the research for this paper.

References

1. Wu SS (2014) Research on China legal system for maritime investigation. *J Sci Commun* 98:51–81
2. Kong J, Fu YH (2017) Comparison study on maritime investigation system between China and Korea. *China waterway transportation*. *Commun* 17:26–31
3. Wang SW (2016) Comparison study on transportation investigation mechanism between China and developed countries. *J Saf Sci Technol Commun* 12:119–124
4. EMSA (2020) Accident report. <http://www.emsa.europa.eu/search-hidden.html?searchword=accident%20report&ordering=newest&searchphrase=all&limit=20>
5. MAIB (2020) Reporting requirements. <https://www.gov.uk/government/organisations/marine-accident-investigation-branch/about>
6. Japan Transport Safety Board (2020) Marine accident and incident reports. <https://www.mlit.go.jp/jtsb/marrep.html>

The Status Quo of China's Implementation of International Maritime Conventions and Suggestions and Countermeasures for Its Participation in Global Governance



Minglu Ma and Guobo Wang

Abstract As a major shipping country, China has been vigorously engaged in global maritime affairs. Through regional and bilateral and multilateral exchanges and cooperation, China has joined international organizations, followed international conventions and voiced China's opinions globally, thus making great headway in management of global maritime affairs. China Maritime Safety Administration (MSA) of the Ministry of Transport has been increasingly engaged in design and development of the administrative and institutional management framework of international maritime affairs, and made full use of the existing cooperation platforms and technology-based promotion channels to promote the implementation of new codes, standards and technologies. So far, China has made a series of breakthroughs in these regards, especially in its Belt and Road Initiative (BRI), and has increased its impact and appeal in global maritime affairs.

Keywords Implementation of international maritime conventions · Participation in global governance · The belt and road initiative

1 The Status Quo of China's Implementation of International Maritime Conventions

In these years, with a focus on the key national and global issues, such as the internationalization of BeiDou Navigation Satellite System (BDS) and greenhouse gas emission reduction, China has been a vigorous participant in leading working groups, formulating demonstration courses and submitting institutional proposals. In 2017, China submitted 46 proposals to the International Maritime Organization (IMO), 34 of which (74%) were submitted by the MSA. These proposals, involving multiple fields of international maritime affairs, further highlighted China's institutional voice in international maritime affairs. The MSA recommended 15 young

M. Ma (✉) · G. Wang
China Waterborne Transport Research Institute, Beijing 100088, China
e-mail: maminglu@wti.ac.cn

junior professional officers (JPOs) to IMO, and conducted technical cooperation with IMO. Additionally, the MSA has successively hosted regional seminars on “Ship Energy Efficiency and Emission Reduction Technologies and Management”, on the development, revision and use of IMO demonstration courses, and on seafarers’ examination and evaluation, and successfully launched exchange programs for PSC officials between China and ASEAN countries.

2 Major Issues for Implementation of Maritime Conventions

2.1 Trends for Implementation of Maritime Conventions

Currently, the mandatory IMO audit scheme, an important manifestation of the globalization of international public governance, is inevitable during the development of the shipping industry and necessary to promote global maritime safety and marine environmental protection. The mandatory scheme puts forward higher requirements for maritime administration, and sets stricter standards for a wide array of fields, such as the security supervision and inspections in port states, special inspections meeting for vessel traffic services (VTS), implementation of crisis prevention, crew management audit, and enterprise management audit.

2.2 Transformation of Implementation

- (1) The maritime administration’s strategic plan, namely, the maritime 5-year plan and annual objectives and tasks, is not full in line with the requirements stipulated in Article 3 of the implementation codes. The overall strategic planning document (the 13th Five-Year Plan) does not fully illuminate how the MSA can fulfill its obligations and responsibilities stipulated in SOLAS and other international conventions, nor is it disclosed to the public officially.
- (2) In regard to maritime legislation and construction of institutional instruments, there is currently a lack of working mechanism for tracking, task decomposition, responsibility and resource allocation for IMO’s maritime provisions that took effect recently. Meanwhile, enterprises, the crew and relevant stakeholders are unable to obtain the relevant texts in time.
- (3) The safeguard mechanism for implementation of international maritime conventions should be further coordinated and improved, so as to facilitate such works as safety management system audit, accident investigation, additional inspections and investigation of overseas PSC detention.

2.3 Implementation Proposals

- (1) With respect to the number of proposals, China has submitted a growing number of proposals over these years, ranking the 5th among the IMO member states. However, there is a big gap between China and the United States, Britain, Norway or Japan in terms of the number of proposals. During 2006–2016, China, a Class A member of IMO, has submitted many proposals, occupying 8% of the total number of proposals submitted by 10 Class A members. However, the number of proposals submitted by Japan, Britain, the United States and Norway accounted for 70% totally.
- (2) In regard to the types of proposals, the proposals submitted by China in 2016 were mostly information-based and comment-based proposals, and there were fewer report-based proposals. These proposals merely focused on voicing the suggestions and advice or taking a stand for certain aspect(s) of a certain issue, and there was no systematic solution or train of thought for solving a certain issue. Among the four types of proposals, the report-based proposal has the greatest influence, while the comment-based proposal has the least influence.
- (3) In terms of issues, the proposals and issues by developed countries mainly highlight the IMO member state audit scheme (IMSAS), goal-based new ship construction standards, anti-piracy and armed robbery, IMO's International Code for ships operating in Polar Waters (Polar Code), etc., while those by China primarily focus on greenhouse gas emission reduction and energy efficiency in international shipping, fishing boats, harmful aquatic organisms in ballast water, remote identification and tracking of ships, etc. This exhibits that China has been following closely the key issues as a major shipping country.

3 New Situations and Requirements of National Strategies and Maritime Development

3.1 Development Opportunities of BRI

The cooperation initiative of the Belt and Road Initiative (BRI), namely, “New Silk Road Economic Belt” and “The 21st Century Maritime Silk Road”, aims to actively develop economic cooperation partnerships with countries along the route of BRI and jointly build a community of interests, destiny and responsibility with political mutual trust, economic integration and cultural tolerance through fully relying on the existing bilateral and multilateral mechanisms between China and relevant countries, and effective regional cooperation platforms.

- (1) Safety risks are unavoidable during the construction of “The 21st Century Maritime Silk Road”. First, there are the huge scale of shipping economy and the actual demands for maritime security. Second, there are daily activities of maritime navigation and the emergency treatment of maritime accidents. Third,

there are potential challenges in maritime security and maritime governance for global maritime security.

- (2) The feasibility of cooperation provides operational conditions for cooperation. The three basic attributes of maritime issues, such as non-zero-sum, the orientation of mutual assistance and benefits, and the tendency to pay attention to absolute benefits, providing a solid theoretical basis and a full practical possibility for maritime security and international cooperation during the construction of “The 21st Century Maritime Silk Road”.

3.2 Cooperation Opportunities in the Lancang-Mekong River

As the important underpinnings for promoting the construction of BRI, the cooperation opportunities in the Lancang-Mekong river have played an increasingly important role, and the five Mekong countries reveal a strong willingness to cooperate with China.

On March 23, 2016, the first Lancang-Mekong Cooperation (LMC) leaders’ meeting was held, and Sanya Declaration was adopted. After that, the five-year plan of action on Lancang-Mekong Cooperation (2018–2022) was adopted, which aims at contributing to the economic and social development of sub-regional countries, enhancing well-being of the people, narrowing the development gap within the region and building a community of shared future of peace and prosperity among Lancang-Mekong countries.

3.3 Issues for the New Situations

There is a lack of corresponding policies for the MSA to integrate and dock the major national strategic plans, such as BRI. As a result, the MSA’s standing, functions and roles have not been fully and effectively manifested.

However, during LMC, there is insufficient cooperation in maritime security, and Agreement on Commercial Navigation on the Lancang-Mekong River signed by the Governments of the People’s Republic of China, the Lao People’s Democratic Republic, the Union of Myanmar and the Kingdom of Thailand in 2000 and the relevant memorandums of understanding and supplementary regulations signed in 2001 are the only legal documents available. The aim of this agreement is to achieve navigation and improve the technical status and construction capacity of vessels operated by the Contracting Parties on the Lancang-Mekong River. However, the requirements in safety technical clauses are outdated and are not in line with the current requirements for vessel safety supervision. Additionally, there is no standard for vessel types in the navigation waters of Lancang-Mekong River and there are not unified requirements for navigation management. Regarding problematic vessels, there are a lack of effective administrative approaches for maritime supervision.

4 Suggestions on Work

4.1 Improvement of Implementation Abilities

4.1.1 Highlighting Transformation of Conventions for Improvement of Implementation Mechanisms

Currently, China must take note of the prominent issues in the current modes for transformation of IMO Conventions facing China.

- (1) Initiation of special studies and establishment of technical procedures for transformation of IMO conventions. In view of China's current legal and regulatory framework and the status quo of implementation, it is suggested to systematically investigate and explore the establishment of feasible technical procedures for the transformation of IMO conventions, aiming at ensuring the integrity of international conventions and timely and accurate implementation of domestic transformation of international conventions.
- (2) Deepening of communication and establishment of a mechanism for coordination and cooperation in implementation. A coordination mechanism between the Ministry of Transport and relevant ministries and commissions and agencies and institutions should be established. Meanwhile, the departments for international affairs, and regulatory, water transport, maritime, search-and-rescue agencies and institutions under the Ministry should jointly establish a response mechanism according to the contents involved in the implementation.

4.1.2 Promotion of Studies on Implementation Strategies

The studies on China's implementation strategies should be promoted to investigate the feasible ways to determine China's implementation strategies. On this basis, China's implementation strategies should be formulated and defined according to the implementation requirements, or the relevant contents of the implementation strategies should be introduced, and all the requirements in the implementation strategies in the IMO Instruments Implementation Code should be incorporated into the preparation framework for the five-year plan in the five-year plan mode.

4.1.3 Enhancement of Work on Implementation Proposals

International maritime affairs have extensively occurred in multiple fields, such as international shipping, shipbuilding, marine environment and resource protection, import and export trade, and economic development. Facing the new development trends, China should expedite the construction of proposals and comprehensively improve maritime capacity and quality to achieve the internationalization of maritime undertakings.

- (1) Establishment of talent teams. The reserve talent pool for IMO's work and specialized research institutions for international maritime affairs should be established, with a view to cultivating strategic talents with a profound understanding of maritime knowledge, a proficient experience in vessel technologies and navigation techniques, a keen understanding of the international conventions, a higher level of English abilities, and a global vision for international maritime affairs. These talents remain committed to safeguarding China's legitimate maritime rights and interests in the international maritime field.
- (2) Strengthening of prospective studies on implementation. The in-depth studies on the development trends of the formulation and revision of IMO's conventions and codes should be performed, with a focus on IMO's major issues, such as greenhouse gas emission reduction by sea transport, IMSAS, goal-based new ship construction standards, and Polar Code. The moderate advance planning should be performed and timely proposals should be submitted.
- (3) Guiding relevant organizations and institutions. China should effectively integrate the resources of relevant domestic institutions and departments, and vigorously guide and encourage relevant Chinese organizations to engage in IMO's affairs.
- (4) Boosting effective organizational coordination. Regarding the IMO's legislative work, the government maritime authorities only play a guiding and auxiliary role, and the relevant industry associations should play a leading role, namely, the relevant industry associations, stakeholders for IMO legislation, launch some research projects in order to safeguard the interests of their own industries and address emerging issues. These projects are supported by the government departments. Relevant professional and technical institutions (classification societies, research institutions and universities), playing a role of technical support for IMO's legislative work, should undertake research projects launched by industry associations and relevant government departments. According to the research reports of relevant technical institutions, the government maritime administrative department submits the proposals to the IMO's relevant committees after consulting relevant industry associations.

4.2 Deepening of International Cooperation

In the current context of complex international multilateral diplomatic situations, IMO member states unremittingly pursue to be a rule-maker. IMSAS periodically assesses and checks the implementation of IMO member States, so as to identify issues and improve the effects of the conventions' implementation. Under the existing cooperation framework, China should jointly submit proposals with other IMO member states and industry organizations to expound common interests and claims. Meanwhile, under the existing framework of regional multilateral exchanges and cooperation, China should remain to be an active participant in the maritime

activities of fully fledged international organizations, and continue to deepen multi-lateral and bilateral exchanges and cooperation with major international maritime organizations globally.

4.3 Advancement of Innovative Cooperation Mechanisms for the BRI

Based on the existing international cooperation framework, China should avail itself of this opportunity of constructing the BRI to innovate international cooperation mechanisms, expand cooperation modes, and strengthen and deepen maritime security and international cooperation with relevant parties.

4.3.1 Intensification of Maritime Cooperation in Mekong River Basin

By reference to the models of the international river navigation commissions, such as the Central Commission for the Navigation of the Rhine (CCNR), the MSA can play a leading role in setting up the Mekong-Lancang River Basin Commission, aiming at proposing systematic standards and suggestions on waterway planning, ship size, vessels' construction quality, navigation requirements and technical specifications. The Commission is entitled to have such rights as legislative right, supervisory right and judicial right, so that it can formulate and revise relevant laws and regulations on navigation safety, supervise the implementation of laws and regulations, have the right to put forward systematic standards and suggestions on waterway planning, navigation requirements and technical specifications, and engage in the conclusion of relevant international agreements.

Meanwhile, the Agreement on Commercial Navigation on the Lancang-Mekong River should be revised or a new agreement on commercial navigation should be formulated. The port states should be entitled to conduct multiple control measures, such as, such as vessel inspection, delay in vessel schedule, operation restrictions and several other major control measures, such as detention, denial of entry and expulsion from port, so as to enhance the coercive force on foreign vessels. Based on the aforementioned measures, the safety supervision basis for foreign vessels is defined, the inspection is performed using the unified technical standards, the procedures for vessel construction and certification are unified, and the automatic identification system (AIS) and vessel traffic service (VTS) are adopted and popularized.

Consolidation of maritime cooperation among coastal countries along the BRI.

4.3.2 Consolidation of Maritime Cooperation Among Coastal Countries Along the BRI

China, together with the coastal countries along the BRI, should establish the maritime cooperation mechanism, convene the maritime safety conference, and conduct technical exchanges, personnel visits and training, and the signing of memorandums regarding vessel inspection, informationization and intellectualization of maritime transport, and improvement of environmental protection level of shipping, achieve mutual recognitions of implementation and crew certificates, and promote the construction of white list, grey list and blacklist of the BRI Memorandum. This not only realizes mutual recognitions between China and the BRI countries, but also promotes mutual recognitions and cooperation in maritime affairs among coastal countries along the BRI, thereby expanding maritime influence and implementing all-round exchanges and cooperation.

4.3.3 Cooperation in Arctic Navigation

Currently, China's proposal for the Memorandum of Understanding on Sino-Russian Maritime Cooperation in Polar Waters has been recognized by Russia. In the next move, China should bolster cooperation with polar countries and expand cooperation mechanisms through the institutionalized construction.

Acknowledgements I am grateful to my colleagues Xu Liansheng and Chen Fengyun, who helped me with the research for this paper.

References

1. Huang YL, Shi XP (2017) Discussion on the establishment of maritime instruments implementation evaluation of China. *China maritime safety*. *Commun* 12:53–55
2. Liu QC, Ou P (2017) Exploration on establishing a new “maritime implementation management system” model. *China maritime safety*. *Commun* 8:24–27
3. Sha ZR (2016) Future development and associated countermeasures on the convention implementation of China maritime safety administration. *Commun* 12:60–62
4. Yang DL (2020) The influence of IMO mandatory audition on China VTS and associated convention implementation strategy. *China maritime safety*. *Commun* 1:30–32
5. Rui HQ (2018) Research on the strategy adopted by MSA for IMO mandatory audition. *International shipping*. *Commun* 41:1–4

Multi-source Data Congestion Recognition Study Based on Fuzzy Pattern Recognition



Junzhuo Li, Wenyong Li, and Bin Guo

Abstract The traffic state identification algorithm based on a single detector data source is often unable to distinguish successfully due to the low prediction accuracy caused by too long distance of fixed monitors and the lack of data caused by communication failure. In recent years, under the background of abundant multi-source data, it has become an important method to use various traffic data information to improve traffic state identification accuracy. This paper proposes a traffic state identification algorithm based on multi-feature fuzzy state recognition method. The algorithm can give weight to traffic data from different sources according to detector types and detection accuracy of the road section to be detected, and identify the data comprehensively. It overcomes the shortcomings of traffic identification methods with single data source, which has low accuracy and is easy to make mistakes.

Keywords Congestion identification · Multi-source data · Fuzzy pattern recognition · Nearness degree · Traffic simulation

1 Introduction

Traffic congestion and traffic accidents are the main traffic problems in large and medium-sized cities. More and more attention has been paid to traffic safety, transport efficiency, environmental pollution and energy consumption caused by traffic congestion. Early traffic management departments mainly rely on artificial ways to find the occurrence of traffic accidents on the road. With the rapid development of society and economy, the road network is constantly constructing, and the demand for traffic is increasing rapidly, which leads to the rapid increase in traffic accidents. In order to improve road traffic safety and operation efficiency, the methods of monitoring road traffic state have been studied extensively [1, 2].

J. Li (✉) · W. Li · B. Guo

School of Architecture and Transportation Engineering, Guilin University of Electronic Technology, Guilin 541004, Guangxi, China

e-mail: ljzshentan@foxmail.com

Most of the early automatic congestion identification (ACI) algorithms were based on traffic data collected by loop detectors, such as traffic volume, occupancy, speed, etc. and traffic states were identified by macroscopic traffic flow methods, such as the fundamental diagram approach and three-phase traffic flow theory [3, 4], Payne and Tignor (1978) published the California algorithm, Persud et al. (1990) designed the McMaster algorithm based on the cusp catastrophe theory. After nearly 50 years of development, fuzzy theory, statistical analysis, time series analysis, artificial neural network and other methods have been applied to the design of ACI algorithm [5–8]. With the advance of information collection technology and information processing technology, video image technology, GPS information, mobile Internet information is also widely used in ACI algorithm design [9–11]. This paper attempts to construct a multi-source detector ACI algorithm by combining detection data from different sources, based on multi-feature fuzzy pattern recognition, to realize fast and efficient prediction of traffic states.

Due to the different types and installation density of detector facilities on different roads, the types and accuracy of data provided by detectors are unconformity greatly. Therefore, when traffic state identification is carried out by combining data from various sources, it is necessary to set up standard state sets according to data types and accuracy provided by detectors in road sections, and to give weight to judgment results of different data. In this paper, by using the method of fuzzy pattern recognition to identify traffic states, will be to identify the various data is divided into two kinds, one kind is parameters to describe the state of the road traffic flow, reflecting vehicles in the average level, represented by the fuzzy vector, another kind is to describe each vehicle running state information, such as lane changing information of the vehicle, vehicle travel time in a road section, the spacing, etc., represented by the fuzzy subset. By calculating the membership degree of the fuzzy vectors to be identified and nearness degree of the fuzzy subsets to be identified onto the standard fuzzy subsets, and synthesizing the weight of the identification model established, the traffic state of the road section is obtained. According to this process, multi-source of traffic data generated by various traffic detectors can be obtained through Vissim simulation, and the results show that the method can be used to identify the data from different sources effectively and has good accuracy (Fig. 1).

2 Characteristics of Observed Data

The early ACI algorithm mainly took the traffic events as the research target, and most of them were based on the traffic volume, occupancy, speed and other road operation characteristics collected by loop detectors. This paper attempts to build an ACI algorithm based on fuzzy pattern recognition according to detection data from various sources under different traffic conditions. In order to find accurate traffic data with diverse types, Vissim simulation method is used to collect a number of detector data. The introduction to the simulated road section is as given in Fig. 2.

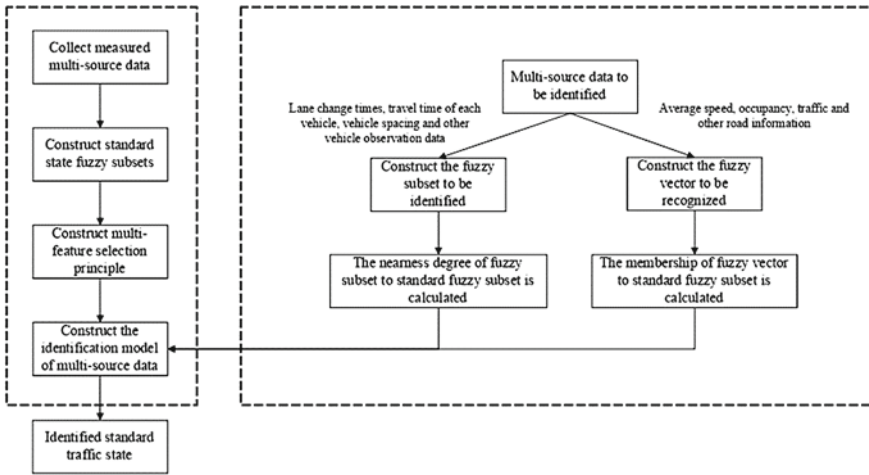


Fig. 1 Flow diagram of the identification model

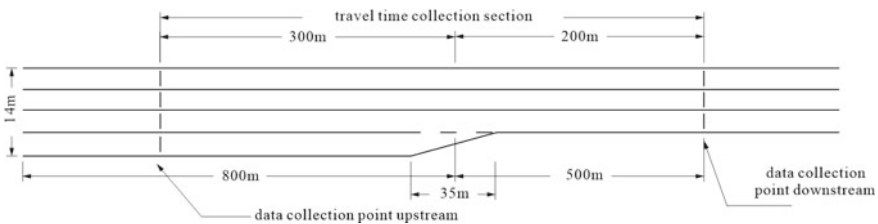


Fig. 2 Introduction of the road section in simulation

The simulation road section has four lanes upstream and three lanes downstream, and travel time collection sections and data collection points set at the designated position of the road section. Except for fixed position detectors, Vissim records all lane changes in the vehicle. Every 10 simulation seconds, Vissim also records the acceleration, speed and other raw data of all vehicles in the road. The simulation lasted for 11,400 s, and the traffic flow gradually increased from 0 to 8000 pcu and then decreased. According to the characteristics of road operation, the traffic state is divided into free flow state, stable flow state, congestion generating state and congestion relieving state. In the free flow state, the driver can drive freely according to his own willing. In the stable flow state, drivers are influenced by other drivers, but they can still freely adjust their driving intention and change lanes. In the congestion generating state, drivers have less opportunity to adjust freely, start queuing and queues are forced to back. After the peak, the queue begins to dissipate, and the congestion continues until the congestion relieving state is over.

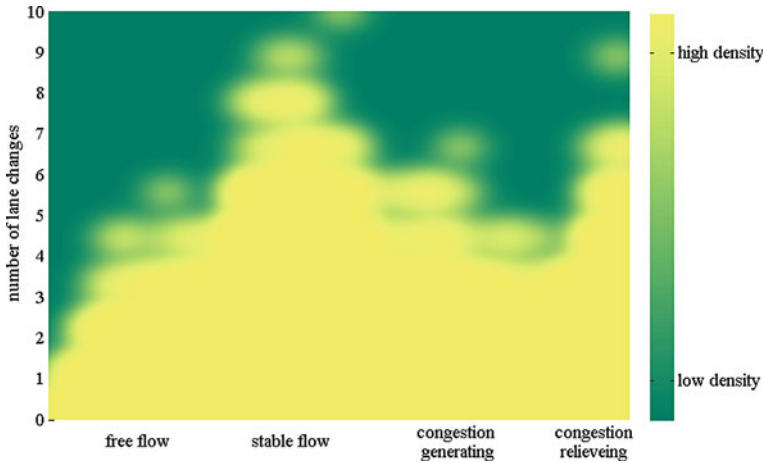


Fig. 3 The distribution of lane changes times in different traffic states

2.1 Number of Lane Changes

In different traffic states, there are obvious differences in drivers' lane change behaviors, among which lane change frequency is the key to observe. Vissim can output the lane changes behavior of all vehicles during simulation and count the number of lane changes. Through the analysis of statistical data, the following conclusions are obtained:

- (1) No matter what kind of traffic conditions, the lane change frequency of the driving vehicle is concentrated on the smaller frequency.
- (2) In the free flow stage, the driver is free to drive the vehicle. Other than the fact that cars in the outermost lane must change lanes when they are in a narrower lane, there is almost no change of lanes.
- (3) During the stage of stable flow, jam gradually arises from vehicles, but there is ample opportunity for overtaking between different vehicles. At this state, traffic lane changes are most frequent.
- (4) During the generating stage of congestion, the self-compression process will increase the density of vehicles, reduce the speed, reduce the chance of overtaking, and reduce the lane change behavior of vehicles.
- (5) At the relieving stage, the effect of vehicles is gradually removed and lane change opportunities are gradually increased (Fig. 3).

2.2 Speed of Data Collection Points

With the change of traffic volume, the distribution of driving speed observed by the upstream and downstream data collection points changes constantly. In the free flow

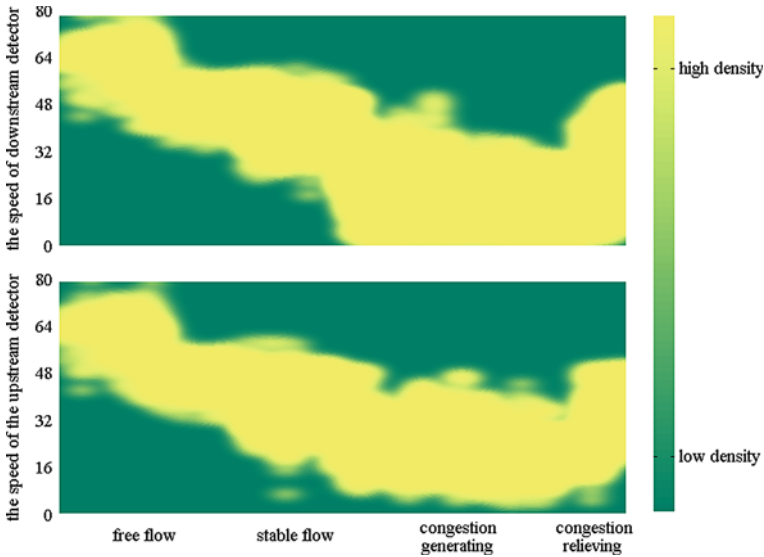


Fig. 4 The distribution of speed in different traffic states

stage, the observed driving speed of upstream and downstream vehicles is almost unchanged with the increase in flow. In the stable flow stage, with the further increase in flow, the observed driving speed of upstream and downstream vehicles decreases rapidly. After the generating of congestion, vehicles will queue up, and the speed decreases slowly. In the upstream part of the bottleneck, some vehicles will slow down and stop. In the downstream part of the bottleneck, the distribution of vehicle speed is relatively concentrated, without stopping (Fig. 4).

2.3 Accelerate Velocity of Vehicles

Drivers expect to drive at a uniform speed on the road, but effect of vehicles leads to the acceleration, and deceleration of the vehicle braking process. Based on the vehicle acceleration per 10 simulation seconds provided by Vissim, the absolute value of each vehicle acceleration are calculated as an average value, which reflects the impact of road vehicles on each other (Fig. 5).

2.4 Travel Time

The travel time distribution is more complicated than other data, and the travel time of the outermost lane is significantly different from that of the inner three lanes. When

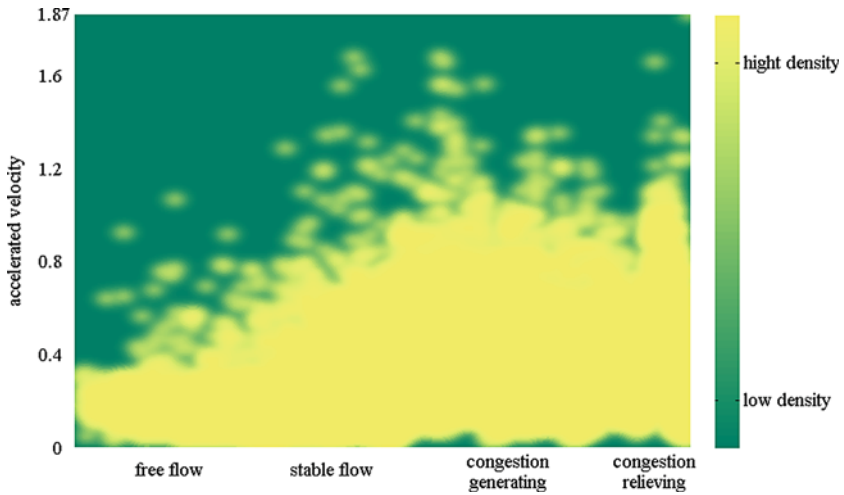


Fig. 5 The distribution of accelerate velocity in different traffic states

congestion occurs, the travel time consumed by vehicles in the outermost lane is significantly higher than that of the other three lanes at the bottleneck section where the four lanes fade to the three lanes. Thus, two different trends in the distribution of travel time are observed (Fig. 6).

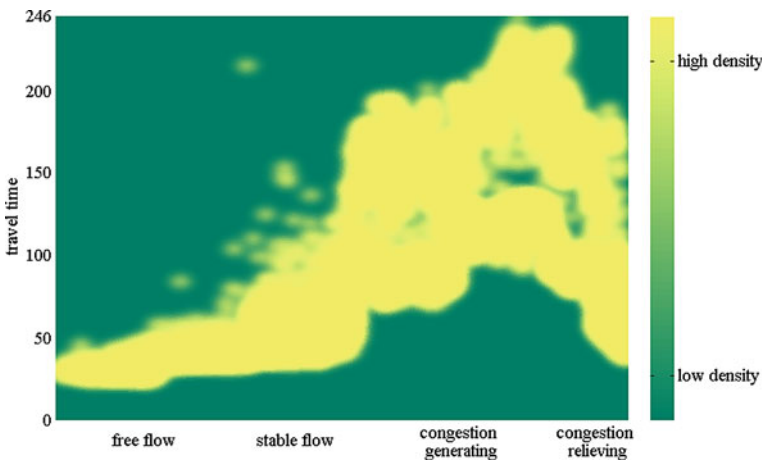


Fig. 6 The distribution of travel time in different traffic states

3 Construct the Membership Function

Different driving behaviors of drivers are found, by observing the changes of various features under different traffic states [12]. When we get the observation data of the road section in a certain period of time, we can identify the current road operation state according to the membership degree of different traffic states to standard states. In this paper, according to the theory of Random Sets, construct the membership function of each feature under different traffic states [13, 14]. The membership function of each traffic state is shown in Fig. 7.

When lane change frequency is small, the membership degree of the four traffic states is relatively high. With the increase of lane changing frequency, the membership functions of different traffic states gradually show differences. Among the four states, the membership function curve of stable flow is obviously different from the other three membership function curves. Consequently, the number of lane changes can be used as a reliability reference for the identification of stable flow state.

The membership function curves of the observed speed of the upstream and downstream detectors have similar shapes. In each traffic state, the four membership function curves have obvious differences, and the speed characteristics have a good performance in the identification of the four states.

Acceleration features have very close membership function under two states of congestion generating and congestion relieving. Therefore, it can only be applied to the identification of free flow, stable flow and congestion, but it cannot effectively identify the generating process and relieving process of congestion (Fig. 8).

The membership function curve of travel time is more complex than other membership function curves. When congestion occurs, the travel time consumed by vehicles in the outermost lane is significantly higher than that of the other three lanes at the bottleneck section where the four lanes fade to the three lanes. Therefore, in the

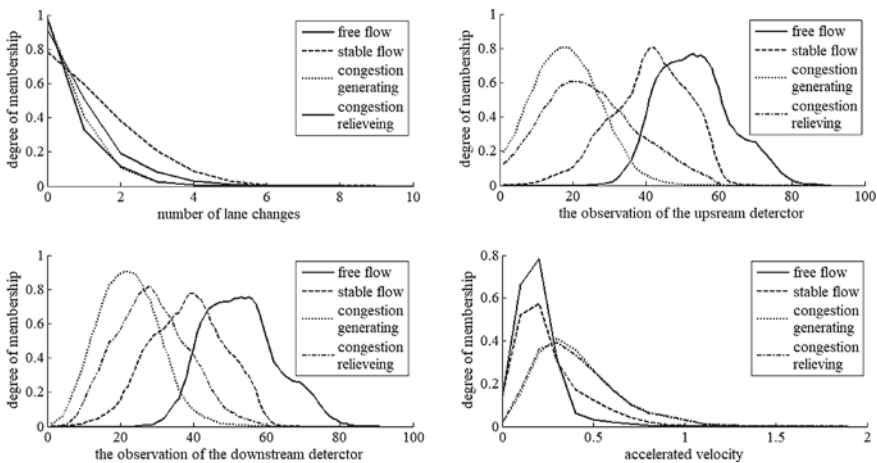


Fig. 7 Membership degree functions of different traffic states

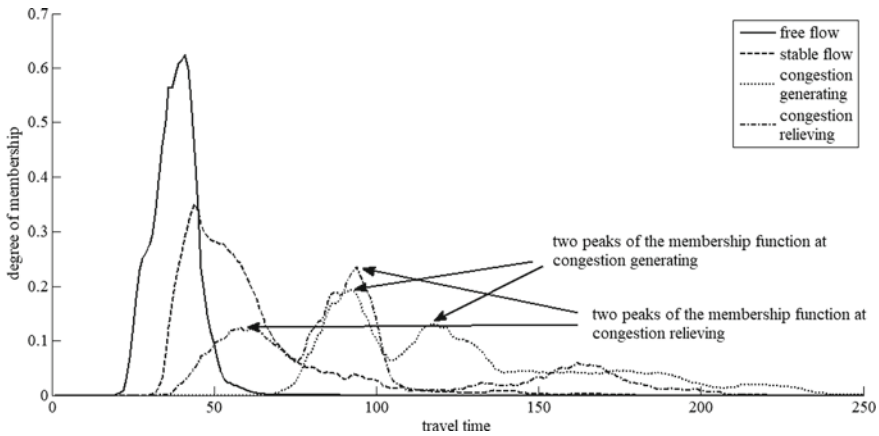


Fig. 8 Membership degree functions of travel time in different states

congestion state, the membership function curves of travel time has two peaks, while in the free flow state and the stable flow state, the travel time has one peak. Due to the different number of peaks in the membership function curve of the congestion state, travel time is very suitable for identifying the congestion state.

4 Construct Multi-feature Fuzzy Model Recognition Model

4.1 Single Feature Selection Principle

Set fuzzy subset \tilde{B} to be identified, and take all vehicles in four states to form four standard fuzzy subsets $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3$, and \tilde{A}_4 respectively. The distance nearness degree is used to represent the approximation degree between membership functions. The distance nearness degree is calculated as Formula (1):

$$\sigma_i(\tilde{A}_i, \tilde{B}) = 1 - \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} |A(x) - B(x)| dx, \quad U = [\alpha, \beta] \tag{1}$$

$U = [\alpha, \beta]$ is the membership function's domain of definition.

When Formula (2) is true, \tilde{B} is closest to \tilde{A}_i , so \tilde{B} belongs to \tilde{A}_i .

$$\sigma_i(A_i, B) = \bigvee_{k=1}^4 \sigma(A_k, B) \tag{2}$$

4.2 Multi-feature Selection Principle

Each fuzzy subset is characterized by 5 features. Therefore, considering the nearness degree of 5 features of the model to be identified to the standard fuzzy subset, it is necessary to construct the principle of contiguity for multiple features.

The contribution of different features to identify different states is different. The larger shapes difference of membership function between different traffic states, the more easily the two states can be identified by these features. When the membership function of a feature in a state is obviously different from the other three standard models, the state identification is more accurate.

The smaller the nearness degree of the state of a certain feature to other states, the greater the contribution to the identification of states. Based on the above analysis, the multi-feature's principle of contiguity is constructed.

Set fuzzy subset $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_4$, and each set is characterized by five features.

$$\tilde{A}_i = (\tilde{A}_{i1}, \tilde{A}_{i2}, \tilde{A}_{i3}, \tilde{A}_{i4}, \tilde{A}_{i5}), \quad i \in \{1, 2, 3, 4\}$$

Fuzzy subset \tilde{B} to be recognized also contains 5 features.

$$\tilde{B} = (\tilde{B}_1, \tilde{B}_2, \tilde{B}_3, \tilde{B}_4, \tilde{B}_5)$$

According to Formula (3), the contribution of the j th feature to the determination of the membership of subset \tilde{B} to the standard subset \tilde{A}_i is calculated.

$$\omega_{ij} = \sum_{k=1, k \neq i}^4 \frac{1}{3a^{\sigma'(A_{ij}, A_{kj})}}, \quad (a > 1) \tag{3}$$

The greater the contribution of a the less sensitive it is to the distance closeness of the standard model. In order to prevent the difference between the values of distance nearness degree of different features, $\sigma'(A_{ij}, A_{kj})$ is the distance nearness degree after standardized according to Formula (4).

$$\sigma'(A_{ij}, A_{kj}) = \frac{6\sigma(A_{ij}, A_{kj}) - \sum_{i \neq k, i < k} (A_{ij}, A_{kj})}{\sum_{i \neq k, i < k} [\sigma(A_{ij} - A_{kj}) - \sum_{i \neq k, i < k} (A_{ij}, A_{kj})]^2} \tag{4}$$

According to the Formula (3), the contribution degree of different features to the judgment of membership degree of different traffic states is calculated. The results are as given in Table 1.

Each feature is weighted according to its contribution to state identification. It is worth noting that the upstream velocity and the downstream observed speed have a very obvious correlation. If the two are weighted as completely independent features, the speed factor will have too much influence on the identification of the model. When

Table 1 Contribution degree of features in different states

Traffic state	Contribution degree of features				
	Accelerate velocity	Upstream velocity	Downstream velocity	Travel time	Number of lane changes
Free flow	1.6003	1.6503	1.866	1.1179	1.2143
Stable flow	0.6729	0.9256	0.8082	0.805	1.9565
Congestion generating	1.132	1.4696	1.4677	1.0079	1.0657
Congestion relieving	1.0271	0.9501	0.9367	0.7661	0.8081

Table 2 Weight of features in different states

Traffic state	Weight of features				
	Accelerate velocity	Upstream velocity	Downstream velocity	Travel time	Number of lane changes
Free flow	0.2812	0.145	0.164	0.1964	0.2134
Stable flow	0.1564	0.1076	0.0939	0.1872	0.4549
Congestion generating	0.2422	0.1572	0.157	0.2156	0.228
Congestion relieving	0.2898	0.134	0.1321	0.2161	0.228

weighting, the contribution of upstream speed and downstream speed can only be calculated by half. The weights of each feature for the judgment of membership degrees of different states are shown in Table 2.

5 Example Verification

In order to verify the identification accuracy of the model, the observation data of 80 vehicles in the congestion generating state were extracted to form a fuzzy subset \tilde{B} to be identified. All vehicles in four states are selected to form four fuzzy subsets $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3$ and \tilde{A}_4 respectively as the standard model. By calculating the nearness degree of the fuzzy set B to the standard fuzzy set, the subordination of the state to be identified to four traffic states can be compared. Firstly, the membership degree function of each feature of fuzzy subset \tilde{B} is calculated. Then, according to Formula (1), the distance nearness degree of the fuzzy subset \tilde{B} 's features to the corresponding features of the standard fuzzy subset is calculated (Fig. 9).

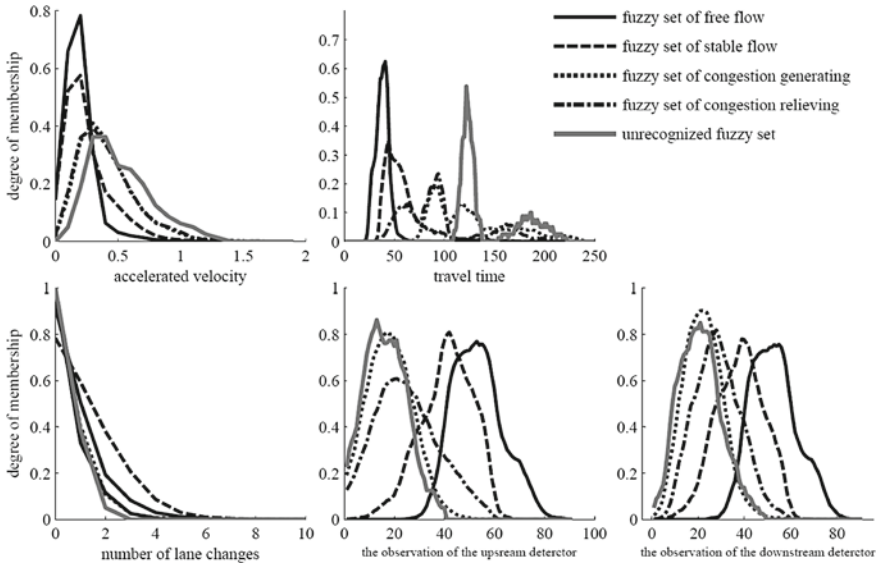


Fig. 9 Membership functions of fuzzy subset to be identified

In order to prevent the difference between the values of distance nearness degree of different features, the membership degree calculated is standardized according to the features, and the standardized values are shown in brackets in Table 3

By calculating the distance nearness degree of the fuzzy subset \tilde{B} to the corresponding features of the standard fuzzy subset, the five features of fuzzy subset \tilde{B} all have the highest nearness degree to each feature of congestion generating state and the fuzzy pattern recognition method can better identify the traffic state.

The standardized values in Table 3 were weighted according to the weights given in Table 2 and then accumulated, the final membership degree of fuzzy subset \tilde{B} to

Table 3 Membership degree of features in different states

Traffic state	Membership degree of features				
	Accelerate velocity	Upstream velocity	Downstream velocity	Travel time	Number of lane changes
Free flow	0.8653 (-1.2164)	0.5869 (-1.2164)	0.6007 (-1.3815)	0.92 (-0.8512)	0.9837 (-1.3376)
Stable flow	0.8972 (-0.5913)	0.6654 (-0.705)	0.7297 (-0.4376)	0.9219 (-0.7184)	0.9009 (-0.5913)
Congestion generating	0.9645 (0.9832)	0.9668 (1.2584)	0.9617 (1.26)	0.9559 (1.6587)	0.9876 (0.9832)
Congestion relieving	0.9629 (0.9458)	0.8754 (0.663)	0.8659 (0.559)	0.9309 (-0.0891)	0.9568 (0.9458)

Table 4 The final membership in different states

Traffic state	Membership degree of feature after weighting					Degree of membership in states
	Accelerate velocity	Upstream velocity	Downstream velocity	Travel time	Number of lane changes	
Free flow	-0.3761	-0.1764	-0.2266	-0.1672	-0.2854	-0.1892
Stable flow	-0.0925	-0.0759	-0.0411	-0.1345	-0.269	-0.2461
Congestion generating	0.2381	0.1978	0.1978	0.3576	0.2242	0.2827
Congestion relieving	0.2741	0.0888	0.0738	-0.0193	0.2156	0.0591

the standard subset is calculated in Table 4. According to the principle of maximum membership degree, fuzzy subset \tilde{B} has the highest membership degree for the congestion generating state, and the state to be identified is the congestion generating state.

In order to judge the accuracy of this recognition method in different traffic states, 200 groups of data were randomly selected from four standard traffic states, and each group of data contained the observed data of 80 vehicles in this state. Due to the observable data of the free flow stage being smaller, in order to test the performance of the method in the small amount of data, 200 sets of data were randomly selected from the free flow state, which only contained the observation data of 40 cars in the free flow state (Table 5).

The results show that the method has a good identification accuracy in all kinds of traffic states, and can also ensure a good identification accuracy in free flow states with a small amount of observed data.

Table 5 Result of identification for each group

Traffic state not identified	Result of identification			
	Free flow	Stable flow	Congestion generating	Congestion relieving
Free flow (40)	200	–	–	–
Free flow (80)	200	–	–	–
Stable flow	–	200	–	–
Congestion generating	–	–	199	1
Congestion relieving	–	–	2	198

6 Conclusion

In order to realize the accurate identification of traffic flow state based on multi-source data, a multi-feature fuzzy recognition algorithm is proposed. Firstly, local observation data are collected to establish fuzzy subsets of standard states. According to the difference of nearness degree of membership function between different standard states, a multi-feature weight model is established. Then, based on the measured data, the fuzzy subset to be identified is established, the closeness degree of the fuzzy subset to be identified and the standard state subset is calculated, and the weight model is applied to the model to obtain the recognition result considering the multi-source monitoring data. The results show that the method can be used to identify the data from different sources effectively and has good accuracy.

Funding This work was supported in National Natural Science Foundation of China through the project Approval No. 61963011.

References

1. Jiang G et al (2002) Development of an automatic incident detection algorithm for freeway based on multi-level alarming system and artificial neural networks. *Int Conf Traffic Transp Stud*
2. Jiang G, Zhang R (2001) Travel time prediction for urban arterial road. *Intell Transp Syst IEEE*
3. Kerner BS (2004) Three-phase traffic theory and highway capacity. *Phys A*
4. Kerner BS (2015) Failure of classical traffic flow theories: a critical review. *E & I Elektrotech Informationstechnik* 132(7):417–433
5. Schönhof M, Helbing D (2007) Empirical features of congested traffic states and their implications for traffic modeling. *Crit Asian Stud* 41(2)
6. Yuan Y et al (2014) Network-wide traffic state estimation using loop detector and floating car data. *ITS J* 18(1–4):41–50
7. Tan H et al (2014) A tensor completion-based traffic state estimation model. In: *Transportation research board annual meeting*
8. Bin R et al (2016) Estimating missing traffic volume using low multilinear rank tensor completion. *J Intell Transp Syst* 20(2)
9. D'Este GM, Zito R, Taylor MAP (2010) Using GPS to measure traffic system performance. *Comput Aided Civ Infrastruct Eng* 14(4):255–265
10. Lui KWK et al (2010) Range-based source localisation with pure reflector in presence of multipath propagation. *Electron Lett*
11. Hawas YE (2007) A fuzzy-based system for incident detection in urban street networks. *Transp Res Part C Emerg Technol* 15(2):69–95
12. Chen D et al (2014) On the periodicity of traffic oscillations and capacity drop: the role of driver characteristics. *Transp Res Part B* 59:117–136
13. Lowen R (1988) Mathematics and fuzziness, part I. *Fuzzy Sets Syst* 27(1):1–3
14. Lowen R (1989) Mathematics and fuzziness, part II. *Fuzzy Sets Syst* 30(1):1–3

Research on Classifiers Used to Identify Dangerous Goods Transportation Vehicles



Haodong Zhang, Qian Cheng, Kuikui Feng, Xiaobei Jiang,
and Wuhong Wang

Abstract With the continuous development of the national economy, the domestic demand for dangerous goods has also increased year by year. Once a traffic accident occurs, it will have a huge impact on the natural environment, road safety, and the safety of people's lives and property. In addition, Advanced Driver Assistance Systems (ADAS) based on sensor technology and advanced control technology provide a good solution for car driving safety. Sensors play a very important role in advanced driver assistance systems. Commonly used sensors mainly include cameras, millimeter wave radars, lidars, etc., which can be used to obtain vehicle internal and external information. This information can help the driver complete the driving task more safely. Therefore, this paper summarizes the current research status of relevant aspects at home and abroad, and compares various vehicle identification and detection algorithms, and uses Haar-features and AdaBoost cascade classifier algorithm to identify dangerous goods transportation vehicles. A total of four classifiers are trained, and the number of positive samples of each classifier is 800, 1200, 1600 and 2000 respectively. Through comparative analysis, it is found that the classifier trained from 1600 positive samples has the best effect.

Keywords Dangerous goods transportation vehicles · Haar-features · AdaBoost cascade classifier

H. Zhang · Q. Cheng · K. Feng · X. Jiang · W. Wang (✉)
Department of Industrial Engineering, Beijing Institute of Technology, Beijing 100081, People's
Republic of China
e-mail: wangwh@bit.edu.cn

H. Zhang
e-mail: ZHDZYW@163.com

Q. Cheng
e-mail: billqiancheng@163.com

K. Feng
e-mail: fengkk1014@163.com

X. Jiang
e-mail: jjiangxiaobei@bit.edu.cn

1 Introduction

There are many uncertain factors in the transportation of dangerous chemicals. According to data released by the U.S. Department of transportation dangerous goods safety bureau. There are about 16,000 to 20,000 traffic accidents related to dangerous goods transportation vehicles every year, accounting for more than 80% of the total number of expressway traffic accidents. It causes economic loss of more than 100 million US dollars every year [1]. Moreover, traffic accidents related to dangerous goods transport vehicles are mostly malignant traffic accidents, which will cause great harm to the natural environment, road safety and people's life and property safety. Due to the characteristics of high center of gravity, heavy load and coupling relationship between tractor and trailer of dangerous goods transport vehicle. These characteristics make the stability limit of dangerous goods transport vehicles low, and it is easy to roll when the steering wheel angle is large, so as to cause huge social property losses [2–4]. Therefore, the ability of dangerous goods transportation vehicles to respond to emergencies is very limited, and other vehicles need to avoid conflicts with dangerous goods transportation vehicles when they encounter dangerous goods transportation vehicles, so as to create a safer road traffic environment for dangerous goods transportation vehicles.

In recent years, with the continuous development of the national economy, the domestic demand for dangerous goods has also increased year by year. As of 2017, my country's total annual transport of hazardous chemicals has exceeded 300 million tons. The transported hazardous chemicals mainly include cyanide, liquid ammonia, liquid chlorine, and oils. Once a traffic accident occurs, it will have a huge impact on the natural environment, road safety, and the safety of people's lives and property. Advanced Driver Assistance Systems (ADAS) based on sensor technology and advanced control technology provide a good solution for car driving safety. Sensors play a very important role in advanced driving assistance systems. Commonly used sensors mainly include cameras, millimeter wave radars, lidars, etc., which can be used to obtain vehicle internal and external information. This information can help the driver complete the driving task more safely. Therefore, this article summarizes the research status of relevant aspects at home and abroad, builds a vehicle detection and recognition algorithm framework, and develops a machine vision-based vehicle identification and early warning system for the transportation of dangerous goods forward.

The current vehicle detection algorithms based on machine vision mainly include methods based on motion information detection, detection methods based on prior knowledge, detection methods based on stereo information, and detection methods based on machine learning [5]. The method based on motion information refers to a method that uses the motion information of the vehicle in the video image collected by the camera to complete vehicle recognition, and the most commonly used method is the optical flow method [6]. The method based on prior knowledge completes the detection of vehicles based on salient vehicle characteristics such as the shadow of the bottom of the vehicle, the characteristics of the edges of the vehicle, the symmetry of

the vehicle, and the characteristics of the taillights [7–9]. The method based on stereo information [10–12] refers to the use of car stereo information to complete vehicle detection. The two most representative methods are based on disparity map method and based on inverse perspective projection transformation method. The detection method based on machine learning [13] uses some feature description operators to extract the feature information of vehicles, and then uses the machine learning method to get the classifier, and finally uses the classifier to complete the vehicle detection.

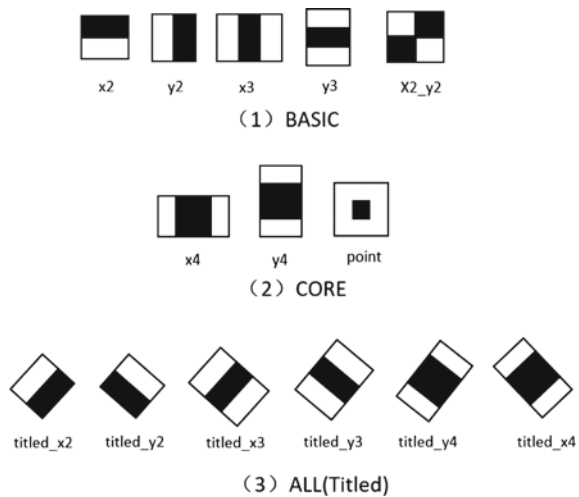
Compared with other detection methods based on machine vision, detection methods based on machine learning are more prominent in recognition performance and robustness [14]. The vehicle recognition algorithm based on machine learning is used more and more times, so this paper decided to use the method based on machine learning to complete the recognition task. Because dangerous goods transportation vehicles have more unique edge features, the Haar-feature can extract the edge features of objects better than other feature description operators. Therefore, this paper chooses an algorithm based on Haar-features and AdaBoost classifier to complete the identification task of dangerous goods transportation vehicles ahead.

2 Method

2.1 Haar-Features and AdaBoost Classifier

The Haar-feature was first proposed by Paul Viola [15], and later extended by Rainer Lienhart et al. to introduce the 45° tilt feature. As shown in Fig. 1, there are 14 Haar

Fig. 1 Haar-features used in OpenCV



features in OpenCV version 2.4.10, which include 5 Basic features, 3 Core features, and 6 Tilted features.

The Haar feature can generate a series of sub-features through magnification and translation in the detection window, but the ratio of the area of the white area to the area of the black area remains unchanged. Take the feature as an example, as shown in Fig. 2, the area ratio of the black area and the white area is always 1:1 during the process of zooming in and panning in the detection window. First, a minimum $\times 2$ feature with a size of 2 pixels is generated at the upper left corner of the red box detection window. Afterwards, a large number of $\times 2$ features with a size of 2 pixels are generated by translation along the x-axis and y-axis in the detection window. The smallest feature is enlarged along the x-axis and y-axis, and then translated along the x-axis and y-axis, and a series of enlarged $\times 2$ features are produced. Then continue zooming and panning according to the above rules until the feature is the same size as the detection window. This produces a complete $\times 2$ series of features.

Since there are relatively few types of dangerous goods transportation vehicles, their shape is also roughly rectangular. Therefore, when identifying dangerous goods transportation vehicles, the edge characteristics are mainly considered. The Basic feature in the Haar feature has a strong ability to describe the edge features of the object. Therefore, the five Basic features of Haar features can be used to detect dangerous goods transport vehicles ahead to achieve better results.

Integrated learning completes the learning task by constructing multiple learners and combining multiple learners together. Therefore, ensemble learning is also called a multi-classifier system. Figure 3 shows the structure of integrated learning. First, generate a series of weak classifiers, and then use a specific strategy to combine them. We can generate weak classifiers from the training data we designed through an existing learning algorithm, such as decision tree algorithms, neural network algorithms, etc. The weak classifiers obtained at this time are homogeneous. If a variety of different learning algorithms are used when generating a weak classifier, such as a decision tree algorithm and a neural network algorithm, the resulting weak

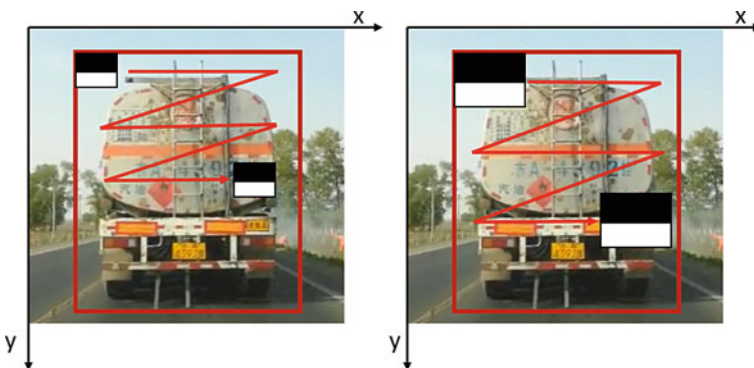
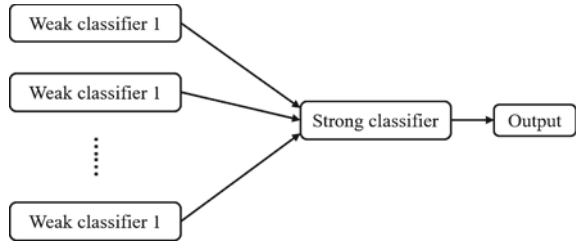


Fig. 2 Schematic diagram of $\times 2$ feature generation sub-features

Fig. 3 Schematic diagram of integrated learning



classifier is heterogeneous. A strong classifier is obtained by integrating multiple weak classifiers in a specific way. Combining multiple weak classifiers through an ensemble learning method can obtain a strong classifier with better generalization ability.

This paper uses the AdaBoost algorithm in the integrated learning algorithm, and its working mechanism is: First, a weak classifier is trained from the training set. Then the weights of the training samples are re-allocated according to the performance of the weak classifier. When assigning weights, the training samples classified incorrectly by the weak classifier get larger weights, and the training samples classified correctly by the weak classifier get smaller weights. Then train the next weak classifier using the training samples obtained with the new weights. Repeat this until the number of weak classifiers reaches the preset value N. Finally, the N weak classifiers are combined according to a certain method.

2.2 Sample Set Design

In this paper, the ensemble learning method in the machine learning method is used to train the classifier used to identify dangerous goods transportation vehicles. The quality of the sample library will directly affect the classification performance of the classifier. The quality of the sample library is mainly reflected in the number and diversity of samples contained in the sample library.

The diversity of samples is mainly affected by the illumination and the resolution of the collection equipment. When the number of samples is too large, it may cause the classifier to take the training samples “too carefully”. As a result, some characteristics of the training sample itself are regarded as the general characteristics of all potential samples, that is, the phenomenon of “overfitting” that is more common in machine learning. Corresponding to it is the phenomenon of “underfitting”, that is, the number of samples collected is too small. The classifier fails to grasp the general nature of the training samples. It is difficult to achieve good generalization ability.

This article collected 3000 positive samples, that is, 3000 pictures of dangerous goods transportation vehicles. In theory, the larger the number of negative samples, the stronger the classifier’s ability to exclude non-dangerous goods transportation vehicles. But in fact, the number of effective negative samples that can be obtained

is limited. And as the number of negative samples increases, the impact of positive samples on the classifier will become smaller and smaller, which will have a great impact on the classification performance of the classifier. Through past experience, it is more reasonable when the ratio of positive and negative samples is between 1:2 and 1:5. Therefore, a total of 9000 negative samples were collected in this article.

Positive sample collection: The positive samples of some dangerous goods transportation vehicles collected are shown in Fig. 4. It is necessary to perform a series of preprocessing on the collected positive samples. First, crop each positive sample, and the result after cropping is shown in Fig. 5. After the cropping process, on the one hand, the classification performance of the classifier can be improved, on the other hand, the rectangular detection frame generated in the detection process can be more accurately circled the target object. Because the following ranging algorithm needs to use the pixel coordinates of the bottom midpoint of the rectangular detection frame. Therefore, the rectangular detection frame more accurately encircles the target object can improve the accuracy of the ranging algorithm. In addition, it is necessary to normalize and grayscale the positive samples. The size of the positive sample used in training the classifier in this article is 24×24 pixels.

Negative sample collection: Negative samples (shown in Fig. 6) are objects other than dangerous goods transport vehicles. The negative samples collected in the article

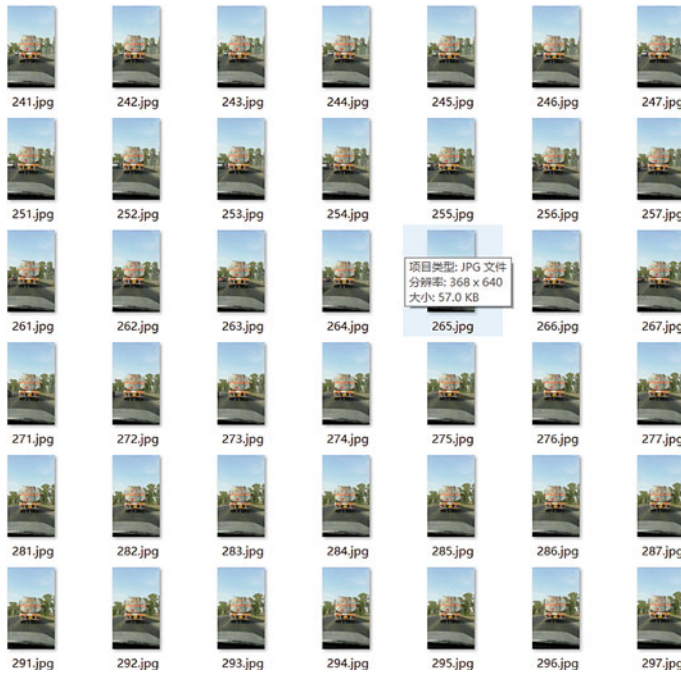


Fig. 4 Original positive sample

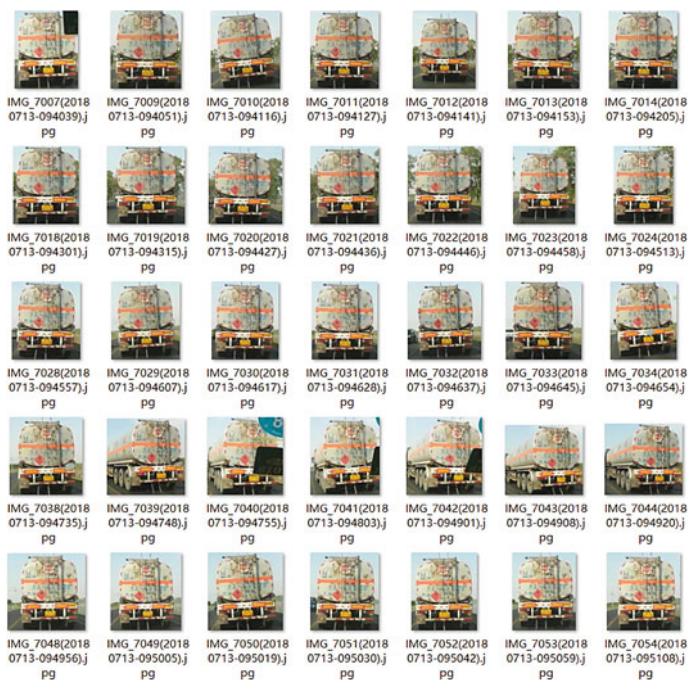


Fig. 5 Positive sample after cropping



Fig. 6 Negative samples

mainly include equipment, buildings, scenery, pedestrians, etc. on the road and its sides, as well as other transportation vehicles except for dangerous goods transportation vehicles. Unlike positive samples, negative samples do not need normalization and grayscale processing, but use the image pyramid and sliding-window to traverse the whole negative sample library, so as to get the negative samples used for training classifiers.

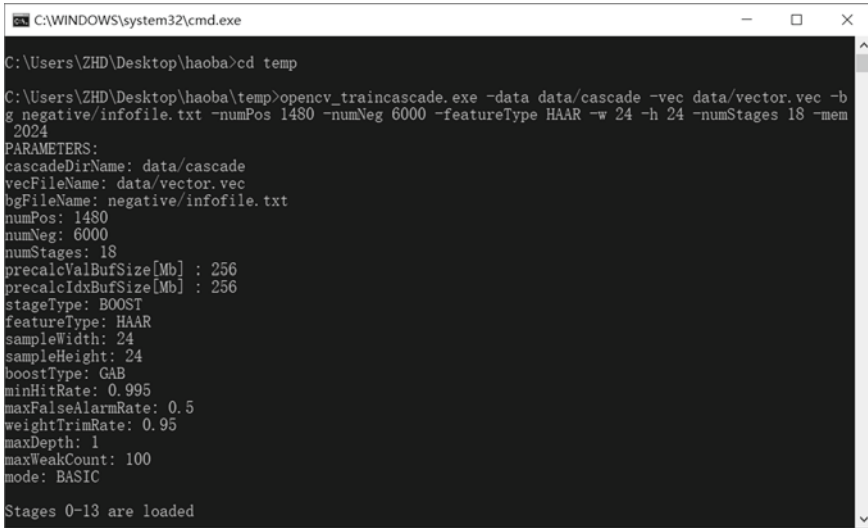
Training classifiers: The image information collected by the camera is affected by the resolution, size and illumination of the camera, which will produce noise which affects the image quality of the camera. In addition, the vehicle is in the process of motion to complete the image acquisition, the captured image will appear fuzzy phenomenon, affect the image quality. Therefore, a series of image preprocessing techniques are needed to reduce the interference information in the image. Therefore, a series of image preprocessing techniques are needed to reduce the interference information in the image. The purpose of image preprocessing is to highlight the effective feature information contained in the image, reduce or eliminate the influence of invalid noise, so as to facilitate the subsequent image processing. In this paper, in order to improve the stability and accuracy of the recognition system of dangerous goods transportation vehicles, the image needs to be grayed and filtered.

In the first two sections, the structure of the classifier and the generation principle of the classifier are analyzed, and the design of the dangerous goods transportation vehicle sample library is completed. This section will train four classifiers with different number of positive and negative samples on the basis of the previous two sections. Through the comparative analysis of the classification performance of the four classifiers, the classifier for dangerous goods transportation vehicles is obtained.

3 Results and Discussion

The acquisition of the cascade classifier based on Haar-features and AdaBoost classifier is done using the 'opencv_traincascade.exe' training program in OpenCV (version 2.4.10), as shown in Fig. 7.

There are 19 parameters to be determined in the training process. Numpos represents the number of positive samples used in each classifier training. Numneg is the number of negative samples used in training each classifier. Numstages represents the maximum number of training classifiers, that is, the number of strong classifiers. featurtype represents the type of feature extraction. Samplewidth represents the width of the sample used to train the classifier. Sampleheight represents the sample height used to train the classifier. boottype represents the AdaBoost algorithm selected. In this paper, gab (gentle AdaBoost) algorithm is selected. maxdepth represents the depth of the weak classifier. maxweakcount represents the number of the largest weak classifiers contained in each step. mode represents the type of Haar feature.



```
C:\WINDOWS\system32\cmd.exe
C:\Users\ZHD\Desktop\haoba>cd temp
C:\Users\ZHD\Desktop\haoba\temp>opencv_traincascade.exe -data data/cascade -vec data/vector.vec -b
g negative/infofile.txt -numPos 1480 -numNeg 6000 -featureType HAAR -w 24 -h 24 -numStages 18 -mem
2024
PARAMETERS:
cascadeDirName: data/cascade
vecFileName: data/vector.vec
bgFileName: negative/infofile.txt
numPos: 1480
numNeg: 6000
numStages: 18
precalcValBufSize[Mb] : 256
precalcIdxBufSize[Mb] : 256
stageType: BOOST
featureType: HAAR
sampleWidth: 24
sampleHeight: 24
boostType: GAB
minHitRate: 0.995
maxFalseAlarmRate: 0.5
weightTrimRate: 0.95
maxDepth: 1
maxWeakCount: 100
mode: BASIC
Stages 0-13 are loaded
```

Fig. 7 The training process of the 'opencv_traincascade.exe' program

In the process of classifier training, the number of positive and negative samples is the most important factor that affects the performance of classifier. In this paper, a number of different classifiers are obtained by a number of different training samples. Then the performance of each classifier is compared and analyzed. Then the best classifier is determined.

There are 3000 positive samples and 9000 negative samples in the sample library. Before classifier training, we need to determine the number of training set and test set, which can be used to evaluate the performance of classifier. If the training set contains most samples and the test set contains few samples, the evaluation results may not be accurate enough. If the test set contains too many samples, the proportion of the training set in the sample library will be reduced. As a result, the model trained by the training set is quite different from the model represented by the sample library, which reduces the “fidelity” of the evaluation results. So far, there is no perfect solution to this problem. Through the past experience, the proportion of training set to sample library is 2/3–4/5, and the rest is the most reasonable as test set. In this paper, four different classifiers are trained, and their parameter settings are shown in Table 1. The training set of the four classifiers contains four times as many samples as the test set. The number of positive samples in training set is 800, 1200, 1500 and 2000 respectively. The number of negative samples is three times that of positive samples. The number of negative samples in the test set is three times that of positive samples.

The other parameters of the four classifiers are consistent in the training process. The maximum number of classifiers is 18. The image feature type is Haar-feature. The sample size is 24 × 24 pixels. The depth of the weak classifier is 1. The maximum number of weak classifications in each order is 100. Basic feature of type in Haar-feature. The remaining parameters can be selected from the default values of the

Table 1 Parameter setting of classifiers

No. of the classifiers	Number of positive samples in training set	Number of negative samples in training set	Number of positive samples in test set	Number of negative samples in test set
Classifier No. 1	800	2400	200	600
Classifier No. 2	1200	3600	300	900
Classifier No. 3	1600	4800	400	1200
Classifier No. 4	2000	6000	500	1500

Table 2 Basic parameters of classifiers

No. of the classifiers	The stages of the classifier	Total number of weak classifiers	Training period
Classifier No. 1	11	55	2 h 22 min
Classifier No. 2	14	83	3 h 18 min
Classifier No. 3	15	106	7 h 41 min
Classifier No. 4	16	142	21 h 32 min

program. After the training, the basic parameters of the four classifiers are shown in Table 2.

As the training progresses, the total false alarm rate of all levels of the classifier will be lower and lower, that is, the probability of a negative sample being incorrectly recognized as a positive sample will be smaller and smaller. On the one hand, it will become very difficult to collect FP using the image pyramid and sliding window method (that is, a negative sample is judged as a positive sample, which is equivalent to a “false detection” or “false alarm”). On the other hand, because the collected FPs are more and more like positive samples, each subsequent level needs to contain more and more weak classifiers. The stages of classifier No. 4 is only one level more than that of classifier No. 3, but the total number of weak classifiers of No. 4 classifier is 36 more than that of classifier No. 3. Moreover, the calculation amount of the No. 4 classifier when searching for FP is much larger than that of the classifier No. 3, which causes the training period of the classifier No. 4 to be nearly 14 h longer than that of the classifier No. 3.

The test set is used to test the classifier’s ability to discriminate new samples. The accuracy of each classifier is shown in Table 3.

As shown in Table 3, compared with the classifier No. 1 and the classifier No. 2, the recognition accuracy difference of the classifier No. 3 and the classifier No. 4 is

Table 3 Accuracy of the classifier

No. of the classifiers	Number of samples in the test set	Number of correctly classified samples	Accuracy (%)
Classifier No. 1	800	504	63
Classifier No. 2	1200	972	81
Classifier No. 3	1600	1472	92
Classifier No. 4	2000	1860	93

smaller and both meet the requirements of accurate recognition. However, considering that the training period of the No. 4 classifier is longer, and the number of stages of the No. 4 classifier and the total number of weak classifiers are larger than that of the classifier No.3, that is, the classifier No. 4 has a larger amount of calculation in the classification process. As a result, the recognition efficiency of No. 4 classifier is not as good as No. 3, so this article chooses classifier No. 3 (that is, a classifier trained from 1600 positive samples and 4800 negative samples) for the detection of dangerous goods transportation vehicles.

4 Conclusion

This paper mainly completed the training of the cascade classifier based on Haar-features and AdaBoost algorithm. First, the classifier algorithm is researched, and the Basic feature in the Haar-feature is determined as the feature description operator of the dangerous goods transport vehicle. Analyze the Haar feature extraction process and the AdaBoost algorithm, and deeply analyze the internal structure and working principle of the classifier. Then, based on the completion of the sample library design, four classifiers with different sample sizes were trained. The positive samples included were 800, 1200, 1600, and 2000, and the number of negative samples included was three times that of the positive samples. By analyzing the internal structure parameters and recognition accuracy of the four classifiers. Finally, it is determined that the classifier trained from 1600 positive samples and 4800 negative samples is used for the identification of dangerous goods transportation vehicles.

References

1. Miao Y, Deng T, Jie F (2012) Application of RFID and GPS technology in transportation vehicles monitoring system for dangerous goods. In: 2012 2nd International conference on remote sensing, environment and transportation engineering (RSETE). IEEE
2. Xiao X, Li H, Liu Y et al (2012) Research on safety monitoring and evaluation system of dangerous goods transportation. IEEE

3. Li S, Tan L, Zhu Y et al (2013) Internet of things for special materials transportation vehicles. In: Green computing & communications. IEEE
4. Chen YG, Wang QJ, Yan LY et al (2011) A study on transport monitoring and emergency rescue management system of dangerous goods. *Appl Mech Mater* 71–78:4091–4094
5. Wei S, Yang L, Chen Z et al (2011) Motion detection based on optical flow and self-adaptive threshold segmentation. *Procedia Eng* 15:3471–3476
6. Park SC, Kim KS, Park KH (2006) Vehicle-monitoring device and method using optical flow. US
7. Bo LI, Zeng ZY, Xiang-Sheng FU (2005) An algorithm for license plate locating based on mathematical morphology and edge characteristics. *TV Eng*
8. Bensrhair A, Bertozzi M, Broggi A et al (2001) A cooperative approach to vision-based vehicle detection. In: ITSC 2001. IEEE Intelligent transportation systems. Proceedings (Cat. No. 01TH8585). IEEE, pp 207–212
9. Handmann U, Kalinke T, Tzomakas C et al (2000) An image processing system for driver assistance. *Image Vis Comput* 18(5):367–376
10. Papageorgiou C, Poggio T (2000) A trainable system for object detection. *Int J Comput Vision* 38(1):15–33
11. Sun Z, Bebis G, Miller R (2005) On-road vehicle detection using evolutionary Gabor filter optimization. *IEEE Trans Intell Transp Syst* 6(2):125–137
12. Niknejad HT, Takeuchi A, Mita S et al (2012) On-road multivehicle tracking using deformable object model and particle filter with improved likelihood estimation. *IEEE Trans Intell Transp Syst* 13(2):748–758
13. Li Z, Zhang K, Chen B et al (2018) Driver identification in intelligent vehicle-systems using machine learning algorithms. *IET Intell Transport Syst* 13(1):40–47
14. Ghahramani Z (2015) Probabilistic machine learning and artificial intelligence. *Nature* 521(7553):452
15. Viola P (2001) Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition*

The Police Perspective on Driving Styles in China: An Interview Study



Siyuan Zhou, Xu Sun, Bingjian Liu, and Gary Burnett

1 Introduction

In recent years, the subject of driving styles has been given increasing levels of attention by researchers around the world (e.g. [23, 34, 38]), because of its close association with road safety issues. It is seen as an important human factor that can potentially explain driver involvement in car accidents and traffic violations, and which may serve as an indicator of driving behaviours among particular groups of drivers [33, 35]. A deep understanding of driving styles enables researchers to gain insights into how people drive and, most importantly, uncover variables that are able to predict their driving behaviours for the ultimate purpose of reducing traffic accidents and further improving road safety [37].

Driving style is defined by Sagberg et al. [29] as “a habitual way of driving, which is characteristic for a driver or a group of drivers” (p. 1251). In line with this definition, a driving style normally refers to a relatively consistent and stable pattern of driving behaviour. Sagberg et al. [29] has also made a clear distinction between the concepts of global and specific driving styles. A global driving style (e.g. reckless driving) is a general concept and comprises a diverse number of specific driving styles with similar underlying motives. By contrast, specific driving styles (such as choice of driving speed, acceleration and safe distancing) are specific habitual driving behaviours held by individuals, and these are seen as the subdomain of a global driving style [9, 13, 29].

Taubman-Ben-Ari et al. [36] suggested four global driving styles, which were later conceptualized in the multidimensional driving style inventory (MDSI), namely: (1)

S. Zhou · X. Sun (✉) · B. Liu
University of Nottingham Ningbo China, Ningbo, China
e-mail: xu.sun@nottingham.edu.cn

G. Burnett
University of Nottingham, Nottingham, UK

reckless and careless; (2) anxious; (3) angry and hostile; (4) patient and careful. The MDSI reflected well the complex and multidimensional nature of driving [37], and has been recognised and validated by many researchers worldwide as an effective tool with which to assess the central domains of driving styles (e.g. [16, 25]). Firstly, the reckless and careless style refers to unsafe and risky driving behaviour that intentionally violates safe driving norms and pursues thrills or excitement during the driving process, but without the necessity to be aggressive or hostile [29, 37]. Secondly, anxious driving relates to “the feelings of alertness and tension, along with ineffective relaxation activities when driving” [19, p. 558]. Thirdly, the angry and hostile style (also known as aggressive driving) is a kind of hazardous driving behaviour that may interfere with and endanger other road users [1]. Finally, patient and careful driving is a well-adjusted and adaptive driving style, including the demonstration of concentration and patience, and compliance with road traffic regulations [37].

National differences in traffic culture have been identified as an important factor that may affect the adopted driving styles [22, 29]. Traffic culture refers to sets of values, norms and attitudes that are used by traffic users in the road traffic environment [27]. This means that the criteria for acceptable driving styles are likely to be determined by the traffic culture of a country, including those associated with government policies, traffic safety regulations and practices [21]. For instance, honking the horn to warn any vehicle of a desire to overtake is not considered a manifestation of aggressive driving in India, but rather an accepted custom, mainly because the lack of rear-view mirrors in many vehicles makes it difficult to monitor traffic on the road behind [10, 29]. Moreover, the perceived safety values held by drivers may also differ due to different cultural values. Drivers in “safe” countries (e.g. the Netherlands) with strict enforcement tend to view ordinary violations as risky and illegal driving behaviour [14]. By contrast, drivers in Turkey are less inclined to see speeding as a serious traffic offense because the speed of traffic on many sections of the road often exceeds the speed limit [22]. Consequently, owing to the fact that driving behaviour can vary significantly between countries, any traffic cultural specificities and social norms, which are likely to govern the driving behaviour informally, should be taken into account before adapting any traffic-related instruments that have been developed elsewhere to a new national context [11].

With regard to the Chinese context, there are several recent studies investigating the traffic culture of China (e.g. [2, 8, 39]). China has a relatively short history of traffic culture in comparison to some Western countries, and only in recent years have traffic laws been amended and implemented to enforce safer driving practice [2]. Drivers, pedestrians and cyclists in China, therefore, are likely to be more tolerant of risk than citizens from other nations, such as the United States and Japan [2]. For instance, research by Shi et al. [31] showed that Chinese drivers were found to engage more in scrambling behaviours (e.g. making turns that fail to give priority to pedestrians or cyclists on a crosswalk) to gain a right-of-way advantage while driving, which violates safe driving norms and often results in traffic accidents. Hence, given the unique characteristics of Chinese traffic culture, it may be presumed that some differences in perspective may exist with regard to the driving styles in comparison to other countries. To the best of our knowledge, few studies have investigated the

interpretations and manifestations of global driving styles (as conceptualized in the MDSI) specifically in the traffic environments of China. Thus, there is a need to examine and explore this topic in detail to construct a more accurate and specific image of the habitual driving behaviour of Chinese drivers.

Furthermore, with regard to the methodology used, most studies related to driving style have traditionally relied on quantitative approaches, including self-reported techniques (e.g. [19, 20]), and objective measurements of some kinematic parameters, such as speed and acceleration (e.g. [9, 40]). While quantitative analysis is well suited to identifying and explaining general patterns or relationships among the selected variables, it may fail to capture complex variables with the use of simplified data, or may neglect some vital aspects of driving style that cannot be quantified [28]. Moreover, although the reliability and usefulness of self-reported measures have been demonstrated by many investigators (e.g. [6, 35]), these responses may still retain potential weaknesses, such as recall bias and self-serving bias, which to some extent make the reporting of one's own driving patterns less trustworthy [15]. On this basis, we have sought to use a qualitative approach as a complementary tool to gain deep insights into driving styles. The use of expert interviews is one of the most important qualitative data collection methods [24]. It was highlighted by Bogner et al. [5] that "the interviewees (as experts in the organizational lifeworld) have internal organizational experience and know-how" (p. 11) so that they can provide more qualified and reliable information. In this qualitative study, the experts are police officers who often undertake road traffic safety enforcement, or who investigate accidents. Because these police officers frequently interact with drivers, observe driving behaviour and specialise in the analysis of road accidents, they are highly likely to offer new and in-depth insights into this research field.

Against this background, an interview study with local road traffic police officers was undertaken to explore the descriptions and manifestations of global driving styles within the particular traffic culture of China. The aim here was to cultivate a deeper understanding of how driving styles are perceived and affected by the unique traffic culture in China, and to obtain evidence of Chinese drivers' habitual driving behaviour.

2 Method

2.1 Participants

The expert group consisted of 10 road traffic police officers (9 males and 1 female) from the Traffic Police Brigade of Ningbo Municipal Public Security Bureau in China, ranging in age from 31 to 58 years ($M = 43.3$, $SD = 7.97$), with their work experience varying between 9 and 38 years ($M = 21.7$, $SD = 8.97$) (Table 1). Seven had extensive working experience of dealing with road traffic infractions and accidents, while the remaining three officers were mainly responsible for publicizing

Table 1 Demographic information of participants

Participant No.	Gender	Age	Years of working experience
1	Male	45	25
2	Male	48	27
3	Male	41	20
4	Male	31	9
5	Male	58	38
6	Male	40	19
7	Male	39	20
8	Male	50	32
9	Male	34	12
10	Female	47	15

issues concerning road safety. All participants were given consent forms to read and sign. The study procedures were approved by the Research Ethics Committee at University of Nottingham Ningbo China.

2.2 Data Collection

An initial contact was made with a leading police officer and permission was granted to undertake an interview in the police station. Other participants, available at the time of the interview, were later randomly invited through the snowball sampling technique, including those working on the tasks of traffic control, accident reporting, analysis and traffic education. A total of 10 officers were involved in semi-structured interviews over two days in July 2019. The interviews lasted approximately 36 min on average, ranging from approximately 15–48 min in length. All the interviews were recorded digitally, following the same procedure to guarantee the reliability of the results. Firstly, the study purpose and the topic of driving styles were introduced in broad terms. Next, the interview questions addressed the main demographic information (e.g. age and years of working experience) and the job tasks of police officers. Then, the researchers asked about their expert opinions with regard to the four central domains of global driving styles within the unique traffic culture of China.

2.3 Data Analysis

Prior to the data processing, all the original audio recordings were checked to ensure that potential personal identifiers and unrelated content had been deleted. Then, all the interviews, including the audio stream of both the interviewers and the respondents,

were transcribed (in Chinese) and double-checked by comparing the transcripts with the original recording files to ensure accuracy. The transcribed documents were later coded into appropriate nodes within the QSR International NVivo for thematic analysis. Here, the categories were created based on the main themes of global driving styles (i.e. reckless and careless, anxious, angry and hostile, and patient and careful).

3 Results

The interview results showed evidence that all four central dimensions of the global driving styles as conceptualized in the MDSI could be found in China, and they were indeed affected informally by the particular traffic culture of China (Table 2).

3.1 *The Reckless and Careless Driving Style*

Several of the respondents ($n = 5$) reported that people who violated traffic rules deliberately when driving (e.g. speeding and running red lights) accounted for a very small proportion of all drivers in Ningbo city. For instance, the phenomenon of speeding was found more among young male drivers who exhibited a strong interest in modified cars for thrills or sensation seeking.

Drivers want to test their cars' performance and enjoyed the thrill of speeding on wide roads.
[Participant 03, male]

One possible explanation is that, if drivers did drive recklessly or carelessly, they would be punished severely by the law. Thus, the implementation of road traffic safety laws forced drivers to obey the traffic regulations under most circumstances.

A few months ago, several drag-racing drivers were punished (by us) for the crime of dangerous driving. [Participant 01, male]

Since the driving environment in China is very complicated, where drivers may have to share the road with a substantially high number of pedestrians and cyclists, three officers indicated that some bad driving practices that intentionally violated safe driving norms, but which would not be punished by the law, were also considered to be reckless and careless. Such behaviour is more commonly exhibited by drivers in China, and they are often associated with a higher risk of road accidents than those which are obviously contravening road traffic safety laws.

Some drivers are inclined to run an amber light at an intersection without decelerating. If there are pedestrians or cyclists infringing the traffic signals to cross the street at their own convenience, and thereby moving suddenly in front of oncoming vehicles, most drivers would fail to react in a timely manner, thus creating a very risky situation which is likely to result in road accidents. [Participant 02, male]

Table 2 Description of driving styles within the traffic culture of China

Global driving styles	Traffic culture of China	Description and manifestation
1. Reckless and careless	<ul style="list-style-type: none"> • Traffic laws. The implementation of road traffic safety laws forces drivers to obey the regulations under most situations • Complicated driving environment. Drivers have to share roads with a large number of cyclists and pedestrians • Risk attitude. Pedestrians and cyclists seem to be more tolerant of risk on the road and are likely to violate the traffic rules for their convenience 	This refers to unsafe and risky driving behaviour. It includes driving behaviour that infringes traffic rules intentionally and obviously (e.g. speeding and running red lights), and also needs to take into account some bad driving practices that fail deliberately to follow safe driving norms (but which may not be punished by the law), such as approaching an intersection without decelerating and running amber lights
2. Anxious	<ul style="list-style-type: none"> • The setting of the driving test. Most novice drivers are trained and tested in a driving school rather than on real roads 	This style is often associated with the feeling of fears or worries about driving, mainly because of insufficient driving experience and inadequate driving skills. It may include some over-cautious driving behaviour, such as driving at a relatively low speed in line with driving school teaching, as well as maintaining an excessive distance between vehicles
3. Angry and hostile	<ul style="list-style-type: none"> • Road rage behaviour 	This style refers to aggressive and hazardous driving behaviour that may endanger other road users and cause accidents. Drivers using this style are likely to engage in the expression of annoyance by hand gestures or verbal abuse, as well as direct confrontation with other drivers/vehicles (e.g. honking the horn, tailgating and weaving)
4. Patient and careful	<ul style="list-style-type: none"> • Complicated driving environment • Risk attitude 	This style employs well-adjusted and adaptive driving behaviour, mostly in compliance with safe driving norms. Establishing appropriate situation awareness and risk evaluation is important for safe driving

Drawn from my own professional experience as a road traffic police officer, I have found that the majority of accidents are actually attributable to drivers' failure to follow safe driving norms, such as merging or changing lanes without checking first. [Participant 06, male]

3.2 *The Anxious Driving Style*

Since driving style refers to a habitual way of driving [29], the responses with regard to driving behaviours which are exclusively determined by an individual's emotional status (e.g. the temporary anxiety caused by the sense of time urgency) have been excluded from this part of the results. It was recognized by most of the police officers ($n = 8$) that the anxious driving style was associated with the feeling of fears or worries about driving, and this was evident more for women and novice drivers.

I myself was in the anxious driving category. I often worried about many potential accidents that could occur while driving, such as colliding with another car. I had no idea about how to respond properly to a variety of road traffic situations. [Participant 10, female]

Novice drivers are more likely to become anxious, holding the steering wheel too tightly and sitting unnaturally erect while driving. [Participant 04, male]

One respondent also highlighted how the location and procedure for the current Chinese driving test (as one aspect of traffic culture) was closely associated with the increasing prevalence of anxious driving among novice drivers. Beginners are often trained and tested in a driving school rather than on real roads. Hence, they were more likely to be anxious when driving in real traffic environments as a result of insufficient driving experience and inadequate driving skills.

In China, the majority of driving schools only teach novices how to pass the driving test, rather than training them to drive safely in real environments. Some people could simply pass both their theory and road tests, and then receive their driving license, all within the period of approximately one month. However, it is obvious that they do not have sufficient driving skills to cope with different real traffic situations, and thus they tend to be anxious when driving. [Participant 06, male]

Furthermore, a few of the participants ($n = 3$) reported that the feeling of anxiety could result in some over-cautious and fear-oriented behaviour, such as driving at a relatively low speed or maintaining an excessive distance from other vehicles.

Sometimes, even on wide roads with a low traffic density, a certain proportion of female drivers would continue to drive at approximately 40 km/h (perceived as a low speed by the police officer) in compliance with instructions learned from their driving school. [Participant 03, male].

Some less skilled female drivers tend to maintain an excessive distance from the car in front, which could further exacerbate road congestion in urban areas, particularly during peak hours. [Participant 08, male]

3.3 *The Angry and Hostile Driving Style*

The angry and hostile driving style is often known as road rage behaviour in China. The majority of the officers ($n = 7$) reported that this style was increasingly common among Chinese drivers, regardless of gender, creating a higher risk of causing accidents and endangering other road users.

Some drivers are likely to view normal behaviour of other road users as aggressive and hostile. [Participant 05, male]

Road rage behaviour is probably more dangerous than some of the risky or illegal behaviour, the consequences of which may result in severe accidents. [Participant 02, male]

There are two main manifestations of the angry and hostile style. Firstly, drivers tend only to use hand gestures or verbal abuse directed at other road users, which is perceived as a relatively mild level of road rage.

Some individuals appear to be good-tempered in their daily lives, but they may engage in verbal expression of annoyance when encountering unexpected and irritating situations while driving. [Participant 01, male]

Secondly, a number of drivers exhibiting angry and hostile behaviour are likely to be involved in direct confrontation with other drivers/vehicles (e.g. honking the horn, tailgating and weaving), often when they want to drive faster than the conditions permit. This could lead to accidents, and drivers with severe road rage behaviours would be punished strictly.

Peak-hour traffic congestion has become a major problem in most cities in China. Some drivers are likely to weave in and out of the traffic and obstruct the path of others to gain a right-of-way advantage. They will take every opportunity to move forward without considering other drivers/vehicles. [Participant 06, male]

3.4 *The Patient and Careful Driving Style*

Several of the participants ($n = 5$) stated that the patient and careful driving style is mostly in compliance with safe driving norms and is endorsed more by experienced drivers. In addition, experienced drivers tend to adjust their driving behaviour to suit different traffic scenarios.

Typically, skilled drivers operate their vehicles in a more standard manner, such as planning the route ahead and paying attention to constantly changing road situations. [Participant 03, male]

Experienced drivers make instant and unambiguous decisions (such as deciding whether or not to cross an intersection) in advance while driving, according to real-time traffic conditions and based on their previous driving experience. [Participant 06, male]

They have nothing to worry about. [Participant 04, male]

Furthermore, skilled drivers were found to be more likely to generate an appropriate risk evaluation based on abundant direct driving experience, which could discourage them from engage in high-risk driving practices.

With a gradual accumulation of driving experience, people may have higher levels of fear about the possible negative consequences brought by risky driving, when compared with novice drivers, and thus they will pursue a more patient and cautious style, and this is particularly the case for those who have previously experienced or witnessed road accidents. [Participant 06, male]

Furthermore, two participants highlighted the fact that the building of situation awareness is an important constitutional element of the patient and careful style for Chinese drivers, especially when taking into account how pedestrians and cyclists in China are seemingly more tolerant of risk on the roads, and will possibly violate regulations for their own convenience.

For instance, at a four-way intersection, there may be some non-motor vehicles violating the traffic rules and running a red light from another perpendicular road segment. Consequently, it is essential for drivers in China to establish situation awareness when approaching an intersection to ensure driving safety. [Participant 01, male]

4 Discussion

A major contribution of this study is to explore the descriptions and manifestations of the four global driving styles (as conceptualized in the MDSI) within the particular traffic culture of China, such as in terms of the driving environment, the location and procedures of driving tests, and traffic safety laws and practices. The interview findings demonstrate that the four central domains are all commonly seen in China, while some are interpreted more specifically and concretely, as influenced informally by the unique traffic culture of China.

Firstly, it was interesting to note that certain bad driving behaviour intentionally violating safe driving norms intentionally (but not necessarily punished by the law) were also perceived by the participating police officers as reckless and careless. Such driving practices (e.g. running amber lights, approaching an intersection without decelerating) are often given inadequate attention by researchers [18, 37], but they could be both risky and dangerous when taking into account the complicated traffic environment in China. In comparison to some Western countries (e.g. the United States), drivers in China have to share the road with a greater number of pedestrians and cyclists [12, 41]. Furthermore, the interview results provide evidence that vulnerable road users in China tend to be more tolerant of risk and will break traffic rules for their own convenience. Uncertainty about the reactions and trajectories of such road users further exacerbates the complexity of many traffic situations [2, 41]. Consequently, this type of bad driving behaviour may also need to be considered, together with those practices obviously and intentionally against traffic safety laws (e.g. speeding and running red lights), when investigating the reckless and careless driving style. This is consistent with research by McKnight and McKnight [17], indicating that only a quite small minority of non-fatal traffic accidents should be attributed to deliberate risk-taking and thrill-seeking behaviour, while a huge number resulted from failure to follow safe driving procedures.

Secondly, it was found that drivers in the anxious category were often worried about their driving owing to a lack of sufficient road experience and inadequate driving skills. In accordance with the studies by Huang et al. [12] and Shi et al. [31], our findings indicate that problem is associated with the location and procedure of driving tests in China. Most people are trained and complete their driving test at a driving school, rather than in a real road traffic environment [12]. Furthermore, in recent years, training at public driving schools has been less strict and shorter in duration than previously, and people mainly learn how to pass the driving test rather than how to drive safely on the road [12, 31]. This might help to explain why some drivers (particularly novices) tend to follow driving norms as instructed at their driving schools, and subsequently engaged in fear-oriented or over-cautious driving behaviour.

Thirdly, with regard to the angry and hostile driving, expressing annoyance through the use of hand gestures and verbal abuse, as well as confronting other vehicles and drivers directly, are regarded as the two main manifestations of this style. The results are in accordance with previous studies relating to aggressive driving [1, 32]. It is worth noting that those people who have adopted this driving style in China are largely driven by the desire to drive faster than the conditions allow, and to acquire a right-of-way advantage.

Finally, following previous studies [4, 26], the establishment of situation awareness and appropriate risk perception are considered by local police officers to be important with regard to the patient and careful driving style. On the one hand, constructing situation awareness allows drivers to seek and gather important and relevant information actively while driving [3, 30], and drivers are able to perceive and interpret any relevant elements of real-time traffic situations for better planning and to control their driving behaviour [3]. On the other hand, the formation of appropriate risk perception can help drivers to choose an effective risk-coping strategy and thereby avoid involvement in high-risk driving activities [7, 26]. Furthermore, unlike some Western countries in which roads are less crowded and where fewer unexpected situations might occur involving pedestrians and cyclists [12], drivers in China tend to have to face more uncertainty about traffic conditions and are, therefore, required to have enhanced driving skills and abilities to suit the various traffic situations for safe driving.

5 Implication

The results of this study have provided some implications for driving styles and road safety. Firstly, the findings offer valuable empirical evidence for the adaptation of traffic-related instruments that are being developed elsewhere into a particular Chinese context (e.g. for creating a Chinese version of the MDSI). It can help to identify a Chinese driver style profile more accurately, by assessing and matching his/her driving behaviours within the four broad domains. Secondly, employing such information in private and public settings (e.g. either by instructors for private

driving lessons, or by safety-officers for the improvement of the traffic culture) may potentially allow drivers to be better aware of their own and other individuals' driving styles, thus encouraging themselves to engage in safer driving practices [37]. For instance, many drivers who exhibit bad driving behaviours (e.g. scrambling behaviours to obtain a right-of-way advantage) do not actually see their actions as reckless and careless [31]. It also suggests the need to take appropriate measures, including the improvement in training programme and safety education, for guiding drivers to operate their vehicles in a safe and careful manner.

6 Conclusion and Limitations

By introducing the expert knowledge of local road traffic police officers through the interview technique, this study has explored the descriptions and manifestations of global driving styles as they pertain to the particular road traffic environments of China. It was found that the central dimensions of driving styles are influenced informally by the traffic culture of China (including the traffic laws, practices and attitudes to risk) and could be interpreted more specifically and accurately. Furthermore, the views of the police officers were found to be much in line with previous scientific research, and these could help us to better understand the driving styles adopted by Chinese drivers, with the support of valuable empirical evidence. However, certain limitations of the study should be noted. Firstly, the police officers' perspectives on driving styles are likely to be geographically focused, meaning that road traffic police officers in other regions of China may have diverging views on each global driving style, or driving behaviour may be different or perceived differently [42]. Therefore, more research data is needed from other regions of China to reflect a broader police view and to achieve generalization of the results. Secondly, in spite of the fact that the interpretations of driving styles are supported by the existing literature, they should still be validated by future experiments, which could provide further support for the understanding of driving behaviour within the particular traffic culture of China.

References

1. Alonso F, Esteban C, Montoro L, Serge A (2019) Conceptualization of aggressive driving behaviors through a perception of aggressive driving scale (PAD). *Trans Res F Traffic Psychol Behav* 60:415–426
2. Atchley P, Shi J, Yamamoto T (2014) Cultural foundations of safety culture: a comparison of traffic safety culture in China, Japan and the United States. *Transp Res F Traffic Psychol Behav* 26:317–325
3. Baumann M, Krems JF (2007) Situation awareness and driving: a cognitive model. In: Cacciabue PC (ed) *Modelling driver behaviour in automotive environments: critical issues in driver interactions with intelligent transport systems*. Springer, London, pp 253–265
4. Beede KE, Kass SJ (2006) Engrossed in conversation: the impact of cell phones on simulated driving performance. *Accid Anal Prev* 38(2):415–421

5. Bogner A, Littig B, Menz W (2009) Introduction: expert interviews—an introduction to a new methodological debate. In: Bogner A, Littig B, Menz W (eds) *Interviewing experts*. Palgrave Macmillan UK, London, pp 1–13
6. Boufous S, Ivers R, Senserrick T, Stevenson M, Norton R, Williamson A (2010) Accuracy of self-report of on-road crashes and traffic offences in a cohort of young drivers: the DRIVE study. *Inj Prev* 16(4):275
7. Carter PM, Bingham CR, Zakrajsek JS, Shope JT, Sayer TB (2014) Social norms and risk perception: predictors of distracted driving behavior among novice adolescent drivers. *J Adolesc Health* 54(5, Supplement):S32–S41
8. Chu W, Wu C, Atombo C, Zhang H, Özkan T (2019) Traffic climate, driver behaviour, and accidents involvement in China. *Accid Anal Prev* 122:119–126
9. Eboli L, Mazzulla G, Pungillo G (2017) How to define the accident risk level of car drivers by combining objective and subjective measures of driving style. *Transp Res F Traffic Psychol Behav* 49:29–38
10. Edensor T (2004) Automobility and national identity: representation, geography and driving practice. *Theory Cult Soc* 21(4–5):101–120
11. Holman AC, Havârneau CE (2015) The Romanian version of the multidimensional driving style inventory: psychometric properties and cultural specificities. *Transp Res F Traffic Psychol Behav* 35:45–59
12. Huang Y-H, Zhang W, Roetting M, Melton D (2006) Experiences from dual-country drivers: driving safely in China and the US. *Saf Sci* 44(9):785–795
13. Itkonen TH, Pekkanen J, Lappi O, Kosonen I, Luttinen T, Summala H (2017) Trade-off between jerk and time headway as an indicator of driving style. *PLOS ONE* 12(10):e0185856
14. Jie L, Van Zuylen HJ, Chen Y, Lu R (2012) Comparison of driver behaviour and saturation flow in China and the Netherlands. *IET Intell Transp Syst* 6(3):318–327
15. Kaye S-A, Lewis I, Freeman J (2018) Comparison of self-report and objective measures of driving behavior and road safety: a systematic review. *J Saf Res* 65:141–151
16. Long S, Ruosong C (2019) Reliability and validity of the multidimensional driving style inventory in Chinese drivers. *Traffic Inj Prev* 20(2):152–157
17. Mcknight AJ, Mcknight AS (2003) Young novice drivers: careless or clueless? *Accid Anal Prev* 35(6):921–925
18. McNally B, Bradley GL (2014) Re-conceptualising the reckless driving behaviour of young drivers. *Accid Anal Prev* 70:245–257
19. Miller G, Taubman-Ben-Ari O (2010) Driving styles among young novice drivers—the contribution of parental driving styles and personal characteristics. *Accid Anal Prev* 42(2):558–570
20. Møller M, Haustein S (2013) Keep on cruising: changes in lifestyle and driving style among male drivers between the age of 18 and 23. *Transport Res F Traffic Psychol Behav* 20:59–69
21. Özkan T, Lajunen T (2011) Chapter 14—Person and environment: traffic culture. In: Porter BE (ed) *Handbook of traffic psychology*. Academic Press, San Diego, pp 179–192
22. Özkan T, Lajunen T, Chliaoutakis JE, Parker D, Summala H (2006) Cross-cultural differences in driving behaviours: a comparison of six countries. *Transport Res F Traffic Psychol Behav* 9(3):227–242
23. Padilla J-L, Castro C, Doncel P, Taubman-Ben-Ari O (2020) Adaptation of the multidimensional driving styles inventory for Spanish drivers: convergent and predictive validity evidence for detecting safe and unsafe driving styles. *Accid Anal Prev* 136:105413
24. Paramonovs S, Ijevleva K (2016) Expert in-depth interview on development and efficiency of “RIGA International Airport” reflecting factors. *Econ Bus* 28(1):98
25. Poó FM, Ledesma RD (2013) A study on the relationship between personality and driving styles. *Traffic Inj Prev* 14(4):346–352
26. Poortvliet PM, Lokhorst AM (2016) The key role of experiential uncertainty when dealing with risks: its relationships with demand for regulation and institutional trust. *Risk Anal* 36(8):1615–1629
27. Pourshams M, Janfada H (2017) Investigation of the effects of religious beliefs on compliance with traffic laws and regulations on Yazd Province drivers (structural equation modeling approach). *J Community Health Res* 6(4):207–215

28. Ragin CC, Amoroso LM (2011) *Constructing social research: the unity and diversity of method*. Pine Forge Press
29. Sagberg F, Selpi, Bianchi Piccinini GF, Engström J (2015) A review of research on driving styles and road safety. *Hum Factors* 57(7):1248–1275
30. Schömig N, Metz B (2013) Three levels of situation awareness in driving with secondary tasks. *Saf Sci* 56:44–51
31. Shi J, Bai Y, Tao L, Atchley P (2011) A model of Beijing drivers' scrambling behaviors. *Accid Anal Prev* 43(4):1540–1546
32. Shinar D (2017) *Traffic safety and human behavior*. Emerald Publishing Limited
33. Sun L, Chang R (2019) Reliability and validity of the multidimensional driving style inventory in Chinese drivers. *Traffic Inj Prev* 20(2):152–157
34. Taubman-Ben-Ari O (2018) What potential role do siblings play in young drivers' driving styles? *Transport Res Part F Traffic Psychol Behav* 58:19–24
35. Taubman-Ben-Ari O, Eherenfreund-Hager A, Prato CG (2016) The value of self-report measures as indicators of driving behaviors among young drivers. *Transport Res Part F Traffic Psychol Behav* 39:33–42
36. Taubman-Ben-Ari O, Mikulincer M, Gillath O (2004) The multidimensional driving style inventory—scale construct and validation. *Accid Anal Prev* 36(3):323–332
37. Taubman-Ben-Ari O, Skvirsky V (2016) The multidimensional driving style inventory a decade later: review of the literature and re-evaluation of the scale. *Accid Anal Prev* 93:179–188
38. Useche SA, Cendales B, Alonso F, Pastor JC, Montoro L (2019) Validation of the multidimensional driving style inventory (MDSI) in professional drivers: how does it work in transportation workers? *Transport Res F Traffic Psychol Behav* 67:155–163
39. Üzümcüoğlu Y, Özkan T, Wu C, Zhang H (2019) How drivers perceive traffic? How they behave in traffic of Turkey and China? *Transport Res F Traffic Psychol Behav* 64:463–471
40. Van Huysduynen HH, Terken J, Eggen B (2018) The relation between self-reported driving style and driving behaviour. A simulator study. *Transport Res F Traffic Psychol Behav* 56:245–255
41. Wang P, Rau P-LP, Salvendy G (2010) Road safety research in China: review and appraisal. *Traffic Inj Prev* 11(4):425–432
42. Zhang Q, Wu C, Chu W, Zhang H (2018) A driver behavior survey study on traffic accidents in China. In: *CICTP 2017*, pp 4530–4540

Simulation Methods of Ultrasonic Guided Wave Propagation Mechanism and Sensor Deployment in Rail



Miao Zhang, Guoqiang Cai, and Ye Zhang

Abstract Ultrasonic guided wave is a recommending tool in the field of structural inspection. With high-speed railway turning into the trend of our national railway, as a direct facility to keep it operation efficient and safe, rail's qualities are guaranteed necessarily. Testing rail with ultrasonic guided waves was public acknowledged. Therefore, to make better use of ultrasonic guided wave for nondestructive testing of rail, verify the feasibility of ultrasonic sensor position. Finite element method was taken to simulate ultrasonic guided wave propagation in head, web and base of rail with the utilization of COMSOL. Determine the central frequency of guide wave excitation via dispersion curve and characteristics of piezoelectric transducer (PZT). On the basis of simulation result the applicable scheme of sensor position in different parts of rail was obtained.

Keywords Ultrasonic guided wave · Rail · COMSOL · Finite element simulation · Frequency selection · Propagation characteristic

1 Introduction

Rails are natural wave guide, testing of rail with ultrasonic guided wave is a promising technology. Some applications of guided waves in rail are investigated by scholars like Sheng [1], Rose et al. [2] and Fan [3]. However, due to the complex section structure of rail, PZT sensors are surface mounted on various places of rail by virtue of experience and in the actual detection of different parts of the rail, results are not satisfactory.

M. Zhang · G. Cai (✉)

State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

e-mail: gqcai@bjtu.edu.cn

Y. Zhang

Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing, China

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

437

W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,

Lecture Notes in Electrical Engineering 775,

https://doi.org/10.1007/978-981-16-5429-9_33

Discussing about guided wave propagation, because of the enormous progress of computer in computational performance and accuracy, finite element method (FEM) technique brings numerical method in investigation the characteristics of wave propagation and scattering. In this paper, COMSOL has contributed to simulate the propagation mechanism of guided waves at base, web and head, the three main detection positions in ultrasonic guided wave rail testing [4]. Therefore, the wave propagation path is displayed.

2 Central Frequency Selection

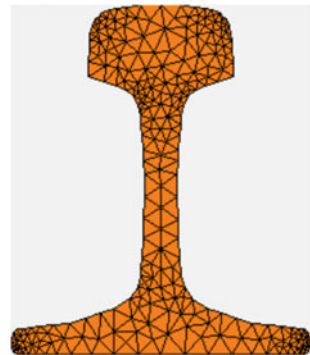
2.1 Dispersion Curve by Semi-analytical Finite Element

Because of the constant cross section of the rail, while calculate the dispersion curve, semi-analytical finite element (SAFE) method can take place of the ordinary finite element method [5], which is accurate but still further complicated facing high frequency and 3D situation.

Discrete the rail cross section into numerous finite small triangle elements like Fig. 1, use u , ε , σ as displacement component, strain field component and stress component of each particle in the rail. Assume X direction of rail be wave propagate direction, and the strain can be expressed by displacement as Eq. (1)

$$\left\{ \begin{array}{l} \varepsilon = [L_x \frac{\partial u}{\partial x} + L_y \frac{\partial u}{\partial y} + L_z \frac{\partial u}{\partial z}] \\ L_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad L_y = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad L_z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array} \right. \quad (1)$$

Fig. 1 Discrete cross section



Discrete the rail cross section into numerous finite small triangle elements like Fig. 1. The displacement of any point in the triangular element is expressed by the shape function as Eq. (2) [6]. Assume the longitudinal displacement field along x-axis of rail is simple harmonic vibration

$$u^{(e)}(x, y, z, t) = Nq^{(e)} e^{i(\xi x - \omega t)} \tag{2}$$

where N is the shape function, $q^{(e)}$ is node displacement vector, ξ is wave number and ω is the frequency. Replace Eq. (2) into Eq. (1), express element strain vector by node displacement as Eq. (3) [7].

$$\varepsilon^{(e)} = [L_x \frac{\partial u}{\partial x} + L_y \frac{\partial u}{\partial y} + L_z \frac{\partial u}{\partial z}] Nq^{(e)} e^{i(\xi x - \omega t)} \tag{3}$$

$$[K_1 + i\xi K_2 + \xi^2 K_3 - \omega^2 M]_M U = 0 \tag{4}$$

by using Hamilton principle [8] the wave equation is derived as Eq. (4), the subscript M is the degree of freedom of the system, K_1, K_2, K_3 are Stiffness matrix, M is mass matrix, U is nodes displacement vector. By solve the wave equation, the dispersion curves of phase velocity and group velocity can be obtained.

A dispersion curves for rail is created like Fig. 2 shows the relevance of phase, group velocity with frequency from 0 to 400 kHz. Generally speaking, higher frequency of the guided wave means shorter wavelength, and more sensitive in damage identification, however rapider attenuation [9]. Consider of reduce wave mode and guarantee long distance propagation, select 0–200 kHz as excite frequency range inspired by general setting of references [1–3, 10].

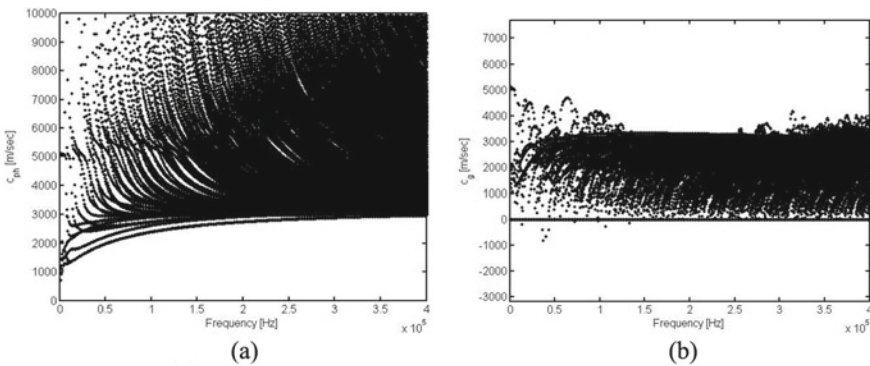


Fig. 2 Phase velocity and group velocity dispersion curves for rail

2.2 Characteristic Analysis of Piezoelectric Transducer

In this paper, piezoelectric ceramic sensor is used as exciter and receptor, we build a simple PZT sensor model which is composed of piezoelectric ceramic and aluminum (The upper part is piezoelectric ceramic lower is aluminum) as universal design, its dimension parameters are shown as Fig. 3.

A piezoelectric sensor has its resonance frequency and frequency-selecting property [11]. Therefore, its working mechanism can be transformed into equivalent circuit like Fig. 4. Where C_0 is the static capacitance of piezoelectric ceramics, C_1 is the dynamic capacitance, R_1 is the dynamic resistance, and L_1 is the dynamic inductance.

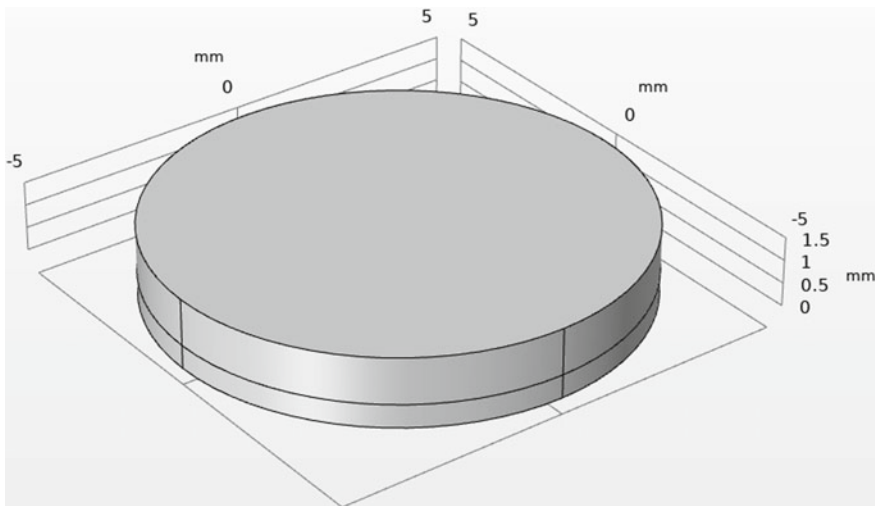
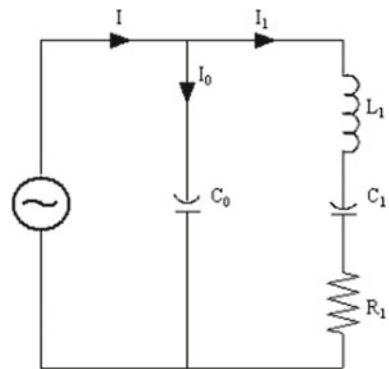


Fig. 3 PZT sensor model

Fig. 4 Equivalent circuit



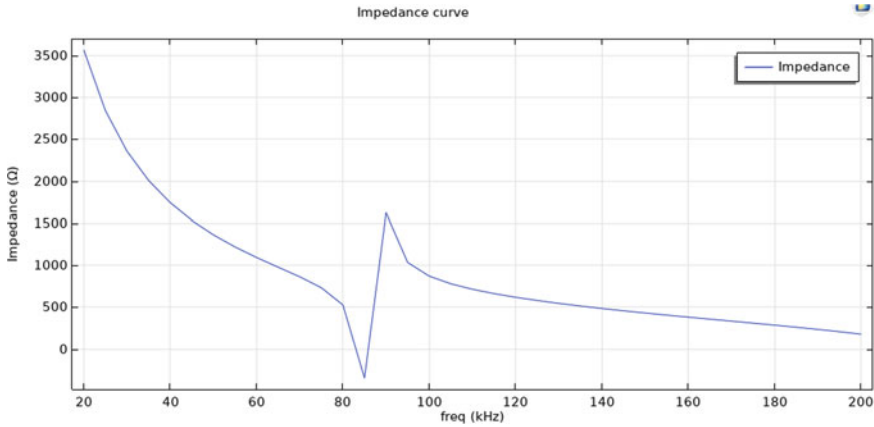


Fig. 5 Impedance curve of PZT sensor

When C_1, L_1 occurs series resonance, the impedance of sensor is in lowest position, emission efficiency reached the maximum value. The expressions of equivalent impedance and admittance are as Eq. (5), where Z is impedance.

$$Z = \frac{U}{I} = \frac{[R_1 + j(\omega L_1 - \frac{1}{\omega C_1})] \frac{1}{j\omega C_0}}{[R_1 + j(\omega L_1 - \frac{1}{\omega C_1})] + \frac{1}{j\omega C_0}} \tag{5}$$

Refer to this frequency-selection property, the sensor’s impedance diagram can be obtained by finite element simulation in Fig. 5.

According to the rail dispersion curve and PZT frequency-selection property, the center frequency of ultrasonic guided wave can be selected as 84 kHz.

3 Modeling

Refer to TB/T 3276-2011 standard [12], 60 kg/m high-speed prefer rail is opted, dimension parameters are shown in Fig. 6a. Build a 500 mm long rail model with material properties shown in Table 1. Set 3 groups PZT sensors separately surface mounted at base, web and head like Fig. 6b.

Set low reflection boundary at end sections, simulate infinite rail.

Five cycles sinusoidal narrow-band wave modulated by Hanning window (Eq. 6) is adopted to be the excitation signal of ultrasonic guided waves with the central frequency is 84 kHz,

$$u(t) = A(t)(1 - \cos 2\pi f_c t/n) \sin(2\pi f_c t + \varphi) \tag{6}$$

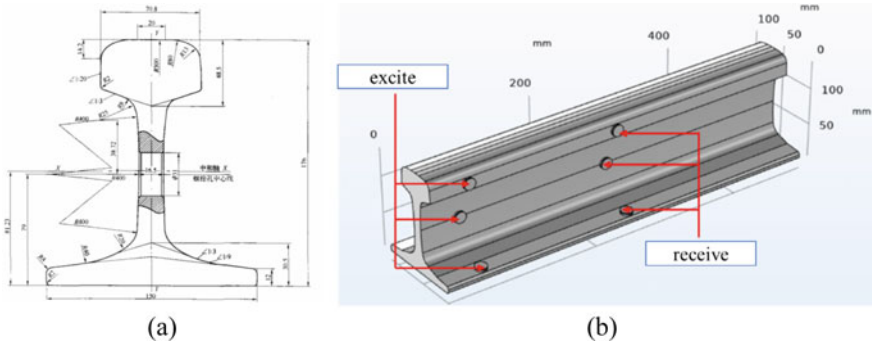


Fig. 6 Dimension parameters of rail (a) and 3D model (b)

Table 1 Rail material properties

	Density (kg/m ³)	Young's modulus (GPa)	Poisson's ratio	S-wave velocity (m/s)	H-wave velocity (m/s)
Value	7850	205	0.3	3172	5934

A(t) is modulation amplitude, f_c is central frequency, φ is the phase and n is wave cycles in this place $n = 5$.

As for the structure meshing, to ensure the calculation accuracy reduce error and avoid unnecessary calculation, generally, each wavelength requires at least 10 nodes [13], refers to Eq. (7) and the maximum size of grid element should be less than 1/10 of excitation wavelength.

$$L_e \leq \frac{\lambda}{10} = cT = \frac{c}{f_c} \tag{7}$$

where λ is wavelength, c is wave velocity, T is cycle and L_e is maximum size of grid element. And simulation time is from 0 to 178 μ s (15 cycles) with step of 0.59 μ s (1/20 cycles). Mesh graph is shown as Fig. 7.

Ultrasonic Guided Waves Propagation in Rail

Generate Hanning window modulated five cycles sinusoidal narrow-band signal in base, web and head PZT sensor separately, with the assistance of displacement image as Fig. 8, propagation path can be observed ocular, therefore received signal and boundary point displacement curve is auxiliary.

Figure 8 shows the displacement caused by excitation at 59, 119, 178 μ s in three positions. After the signal excitation, guided wave occurs after multiple boundary reflections and propagate along the longitudinal direction.

The earliest waves accumulate around every PZT sensor, then reflected by structure boundary, compare with base section more scattering is produced in the head and web of rail where are thicker, and other more complicate waves mode generate by

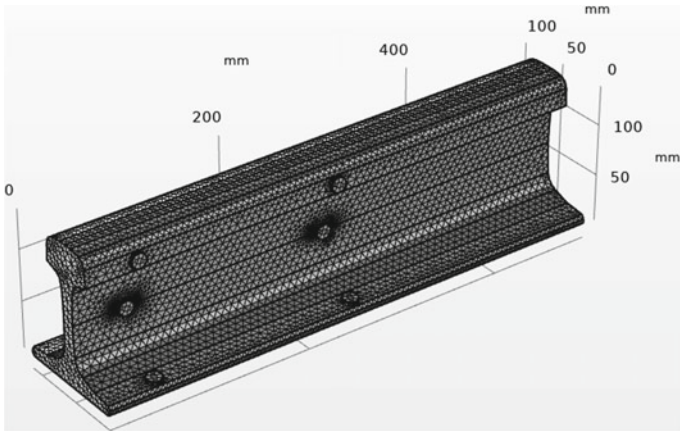


Fig. 7 Rail model mesh

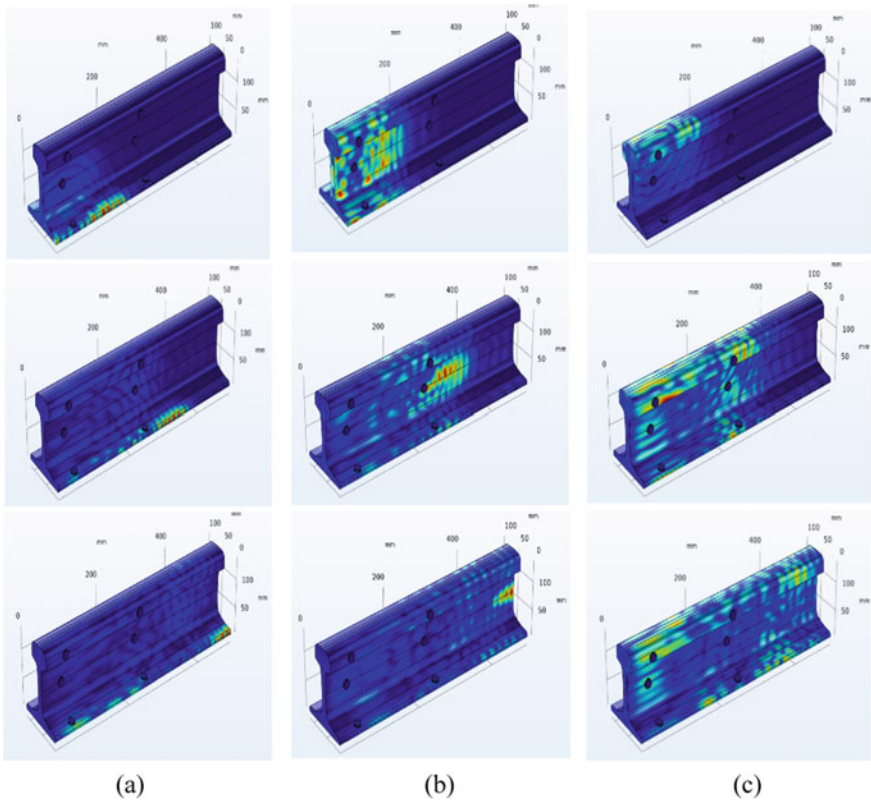


Fig. 8 Displacement image at 59 μs, 119 μs, 178 μs (59 μs, 119 μs, 178 μs from top to bottom) in base (a), web (b), head (c)

reflect wave arrived receive sensor after direct waves, in this case only direct waves are discussed.

Because of distinct section thickness and structure characteristic of rail, guided waves propagate diversely within different section, substantial wave divergence occurs in web and head excitation, big amplitude wave emerges in other area besides excite position with time delay. Figure 9a–c are boundary displacement curves in diverse guided wave excite positions and Fig. 9d–f are the receive.

Combining boundary displacement curves and received signal we can observe that base excite cause slight wave divergence in web and head, as consequence, notable signal got by base receptor but few signals are received by PZT sensor mounted at web and head of rail. Meanwhile web and head excite brings lager wave into other position, impels considerable signal received by every receptor, which make cross-section testing feasible.

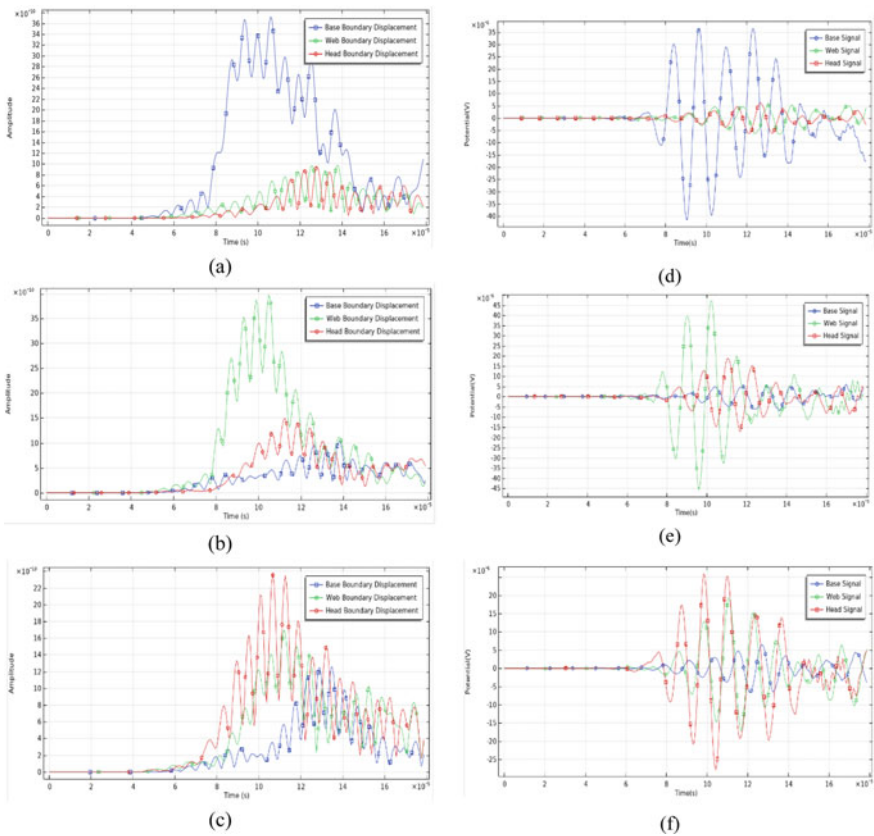


Fig. 9 Base (a), web (b), head (c) boundary displacement curve and received signal in base (d), web (e), head (f) of rail

When testing partial rail with ultrasonic guided waves, as for base detection the good way is excite in base and receive in base and signal originate in head may work as well. For web detection, when excitations from head, web receptor can also accomplish detection besides of web excite web receive mode. And head detection needs to generate waves at head and web.

4 Conclusion

Ultrasonic guided waves are established for long-range inspection, in this paper a 500 mm long 60 kg/m rail model with 6 PZT sensors are established, and with the utilize of COMSOL simulation the ultrasonic guided wave propagation process of rail can be observed ocular.

Combining the dispersion curve of rail and the PZT sensor impedance curve, the appropriate guided wave excitation frequency was selected, and improve the potency of guided wave detection.

On the basis of simulation result, as for different rail section testing, applicable scheme of PZT sensor position in different parts of rail is obtained with the observation of boundary displacement curves and received signal. Cross-section detection methods of rail are proposed.

Acknowledgements The paper is partial support by National Key R&D Program of China (2018YFB1201601-07).

References

1. Sheng H. Study on guided wave detection method of rail bottom oblique crack and weld. Nanchang Aeronautical University, Jiangxi
2. Rose JL, Mudge P, Avioli MJ (2004) Guided wave inspection potential of defects in rail. *NDT & E Int (Independent Nondestructive Testing and Evaluation)* 37(2):153–161
3. Fan Z (2016) Research on guided wave detection technology for switch rail and long center rail. China Academy of Railway Sciences, Beijing
4. Kang Z, Wang M, Chen H (2008) Ultrasonic flaw detection of railway turnout and rail profiled section. *Railway Tech Supervision* 36(6):16–18. <https://doi.org/10.3969/j.issn.1006-9178.2008.06.006>
5. Hayashi T. Guided wave dispersion curves derived with a semi—analytical finite element method and its applications to nondestructive inspection. *Jpn J Appl Phys*
6. Loveday PW (2012) Guided wave inspection and monitoring of railway track. *J Nondestructive Eval* 31(4):303–309
7. Bartoli I, Marzani A et al (2006) Modeling wave propagation in damped waveguides of arbitrary crosssection. *J Sound Vib* 295(3):685–707
8. Xiao K, Song W (2008) Study on two dimensional finite element ultrasonic wave equation. *Foreign Electron Meas Technol* 27(5):9.11
9. Liu Q (2013) Study on propagation characteristics of ultrasonic guided wave in rail. Beijing University of Technology

10. Hayashi T, Miyazaki Y, Murase M et al (2007) Guided wave inspection for bottom edge of rails. AIP Conf Proc
11. Gao T, Zeng J, Peng D (2007) Analysis of broadband matching characteristics of piezoelectric ceramic transducer. Acta Acoust Sin 32(2):97–102
12. TB/T 3276-2011. Rail for high speed railway
13. Yan Y (2015) Numerical simulation and experimental study on flaw detection of switch rail based on ultrasonic guided wave. Jinan University

Passing Ability Impact Factors Analysis of Heavy Tractor



Hao Zhang, Cheng-qiang Zong, Chuan-jin Ou, and Yue Li

Abstract Passing ability of articulated vehicles is analyzed, which are combined of different 6×4 tractors and different semitrailers. According to the requirements of GB 1589-2016 ‘Limits of dimensions, axle load and masses for motor vehicles, trailers and combination vehicles’, passing ability of different structure articulated vehicles are simulated by using AutoTURN software. Tractor front overhang, distance between center of tractor fifth wheel and tractor first axle, semitrailer first wheel space is selected as the main factors of articulated vehicle passing ability, in order to analyzing passing ability impact factor of articulated vehicles type and structure. Simulation analysis is used to study how tractor front overhang, distance between center of tractor fifth wheel and tractor first axle, and semitrailer first wheel space impact passing ability. Some reasonable suggestions is proposed by analysis how parameters change impact passing ability. The validity of the simulation model is validated by the test data of ‘3rd Swap Trailer Transport Recommended Vehicles Test’. The results provide a useful reference for improving articulated vehicle passing ability level and promote the development of vehicle industry technology.

Keywords AutoTURN · Articulated vehicle · Impact factors · Passing ability · Simulation analysis

1 Introduction

GB 1589-2016 ‘Limits of dimensions, axle load and masses for motor vehicles, trailers and combination vehicles’ is the primary national standard which refers to the limits of axle load and total mass in China. This standard is revised based on the original standard GB1589-2004 ‘Limits of dimensions, axle load and masses for road vehicles’. The implement of this standard offered legal basis for road administration,

H. Zhang (✉) · C. Zong · C. Ou · Y. Li

Research Institute of Highway Ministry of Transport, Beijing 100088, China

Key Laboratory of Operation Safety Technology on Transport Vehicles Ministry of Transport, Beijing 100088, China

and played an important role in calming down load over signs, over load and over limits, and also influenced the development of multi-axis heavy-duty special semi-trailer, multi-axis heavy-duty trucks and vehicle lightweight [1]. In this standard, the requirements for vehicle passing ability played an important role for improving vehicle passing ability and travel safety.

AutoTURN is the vehicle swept path analysis software of choice for transportation engineers, architects, and planners worldwide. AutoTURN is used to confidently analyze road and site design projects including intersections, roundabouts, bus terminals, loading bays, parking lots or any on/off-street assignments involving vehicle access checks, clearances, and swept path maneuvers [2–4]. Articulated vehicles' passing ability is analyzed based on AutoTURN according to the requirements of passing ability in GB 1589-2004 'Limits of dimensions, axle load and masses for road vehicles'. Passing ability impact factor of articulated vehicles type and structure are analyzed.

2 Preparation for Simulation

2.1 Simulation Model

Articulated vehicles combined of 6×4 tractor container semitrailer and van semitrailer respectively are used as simulation model. Passing ability of articulated vehicles are analyzed based on those simulation model. Parameters of articulated vehicles are shown in Table 1. Simulation data setting is illustrated in Fig. 1.

Table 1 Parameters of vehicles

Tractor (mm)	Front overhang (L_{Tfo})	Wheel space (L_{Tw})	Fifth wheel lead (L_{Tfl})	Driving cab length (L_{Tc})
6×4 Tractor 1	1410	3200	1109	2500
6×4 Tractor 2	1365	3200	1050	2500
Semitrailer (mm)	Front overhang (L_{Sfo})	Wheel space (L_{Sw})	Trailer length (L_S)	Total length (L)
Container semitrailer	987	7080/1310/1310	12,400	16,290
Van semitrailer	1365	8100/1310/1310	14,600	17,830

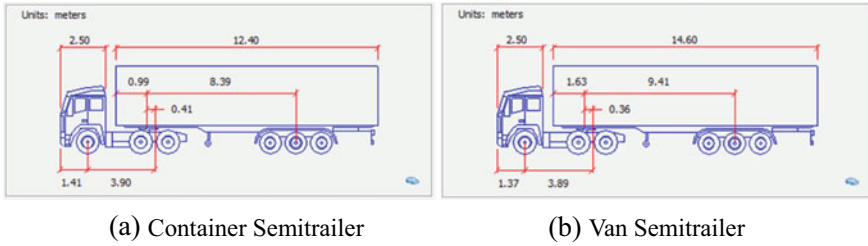
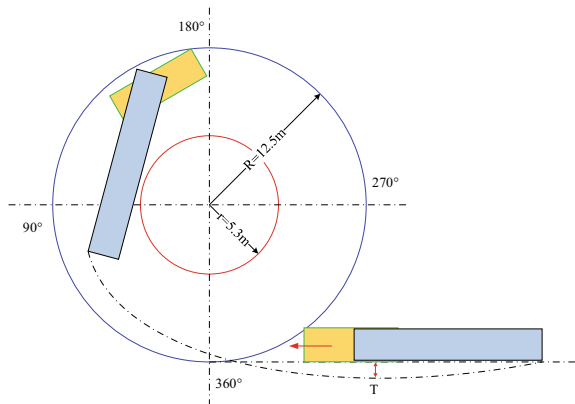


Fig. 1 Simulation data setting

2.2 Simulation Condition Settings

Articulated vehicles' passing ability simulation is based on requirements of passing ability in GB 1589-2016. Appendix B in this standard gives a method of how to test vehicle turning clearance circle diameter and out value method [5, 6]. The inner radius of each angle is measure value of intermediate shaft of Semitrailer track. Tractor front overhang, fifth wheel lead, tractor wheel space (the distance between center of tractor fifth wheel and tractor first axle), and the semitrailer wheel space (the distance between center of semitrailer fifth wheel and semitrailer first axle) is selected as key parameters in passing ability simulation. Tractor front overhang, tractor wheel space, fifth-wheel lead are make 50 mm adjustments separately for comparing impact of articulated vehicle passing ability in different simulation. Tractor wheel space is made 100 mm adjustments for comparing impact of articulated vehicle passing ability in different simulation. The quotient of adjustment difference and simulation result difference of original value is selected as evaluation index (Fig. 2).

Fig. 2 Experimental schematic diagram



3 Simulation Results Analysis

3.1 Container Semitrailer

(1) Tractor Front Overhang

Length of tractor front overhang is 1410 mm, which is changed as previously mentioned only and other parameters remain the same in passing ability simulation. In this condition, changes of tractor front overhang equal to articulated vehicle length changes. The results of simulation are shown in Fig. 3.

When the length of tractor front overhang changes in the range of 200 mm, the simulation results of turning clearance circle at different angles remain the same. In conclusion, adjusting the length of tractor front overhang only seldom affects the passing ability of articulated vehicle.

(2) Fifth Wheel Lead

Length of fifth wheel lead is 1109 mm, which is changed as previously mentioned only and other parameters remain the same in passing ability simulation. In this condition, length of articulated vehicle don't change. The results of simulation are shown in Fig. 4.

When the length of fifth wheel lead changes in the range of 200 mm, the difference of turning clearance circle at different angles of 90°, 180°, 270°, 360° are -10 mm,

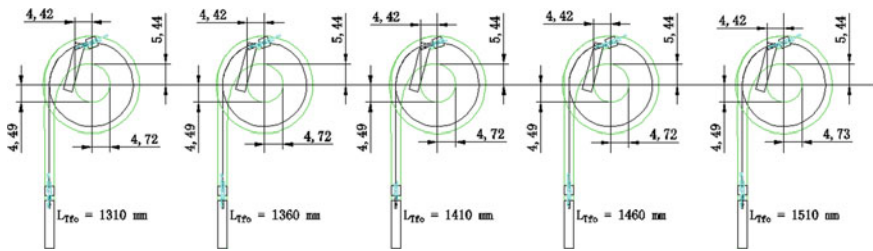


Fig. 3 Simulation results of tractor front overhang change

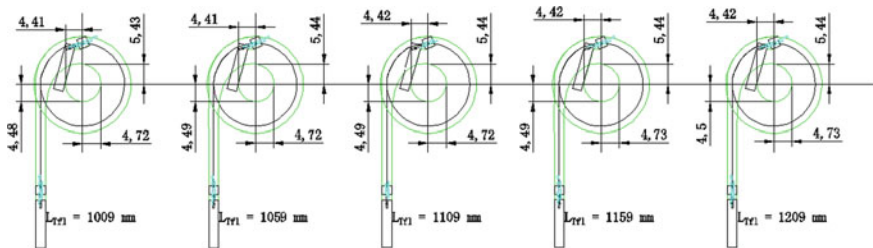


Fig. 4 Simulation results of fifth wheel lead change

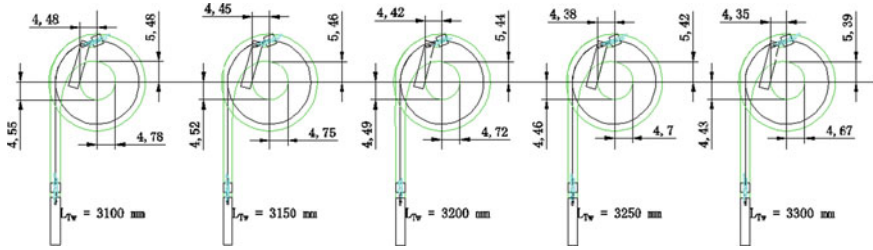


Fig. 5 Simulation results of tractor wheel space change

−10 mm, −20 mm, −10 mm and the quotients of evaluate indexes are −0.05, −0.05, −0.10, −0.05 corresponding. In conclusion, adjusting the length of fifth wheel lead only seldom affects the passing ability of articulated vehicle.

(3) Tractor Wheel Space

Length of tractor wheel space is 3200 mm, which is changed as previously mentioned only and other parameters remain the same in passing ability simulation. In this condition, length of articulated vehicle don't change. The results of simulation are shown in Fig. 5.

When the length of tractor wheel space changes in the range of 200 mm, the difference of turning clearance circle at different angles of 90°, 180°, 270°, 360° are −90 mm, −110 mm, −120 mm, −130 mm and the quotients as evaluate indexes are 0.45, 0.55, 0.60, 0.65 corresponding. In conclusion, decreasing the length of tractor wheel space reasonably will enhance the passing ability of articulated vehicle effectively.

(4) Semitrailer Wheel Space

Length of semitrailer wheel space is 7080 mm, which is changed as previously mentioned only and other parameters remain the same in passing ability simulation. In this condition, length of articulated vehicle don't change. The results of simulation are shown in Fig. 6.

When the length of t semitrailer wheel space changes in the range of 200 mm, the difference of turning clearance circle at different angles of 90°, 180°, 270°, 360° are

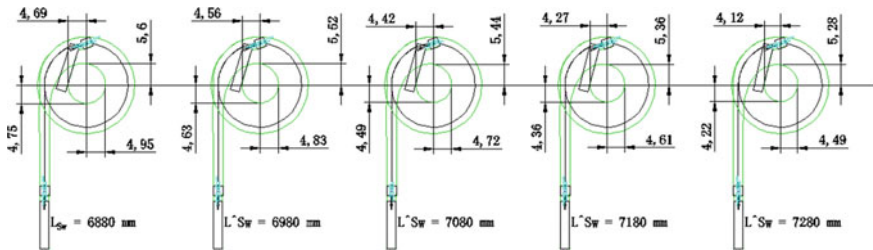


Fig. 6 Simulation results of semitrailer wheel space change

−320 mm, −460 mm, −530 mm, −570 mm and the quotients as evaluate indexes are −0.80, −1.15, −1.33, −1.43 corresponding. In conclusion, decreasing the length of semitrailer wheel space reasonably will enhance the passing ability of articulated vehicle effectively.

In summary, tractor wheel space and semitrailer wheel space are the main impact parameters in passing ability simulation of articulated vehicle, which is combined of 6×4 tractor and container semitrailer. Decreasing the length of tractor wheel space and semitrailer wheel space reasonably will enhance the passing ability of articulated vehicle effectively. In order to improving articulated vehicle passing ability, it is recommended to decrease the length of tractor wheel space 100–150 mm and the length of semitrailer wheel space 150–250 mm.

3.2 Van Semitrailer

(1) Tractor Front Overhang

Length of tractor front overhang is 1365 mm, which is changed as previously mentioned only and other parameters remain the same in passing ability simulation. In this condition, changes of tractor front overhang equal to articulated vehicle length changes. The results of simulation are shown in Fig. 7.

When the length of tractor front overhang changes in the range of 200 mm, the simulation results of turning clearance circle at different angles remain the same. In conclusion, adjusting the length of tractor front overhang only seldom affects the passing ability of articulated vehicle.

(2) Fifth Wheel Lead

Length of fifth wheel lead is 1050 mm, which is changed as previously mentioned only and other parameters remain the same in passing ability simulation. In this condition, length of articulated vehicle don't change. The results of simulation are shown in Fig. 8.

When the length of fifth wheel lead changes in the range of 200 mm, the difference of turning clearance circle at different angles of 90° , 180° , 270° , 360° are −10 mm,

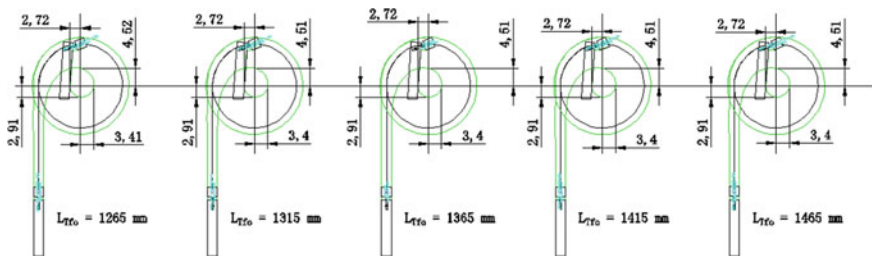


Fig. 7 Simulation results of tractor front overhang change

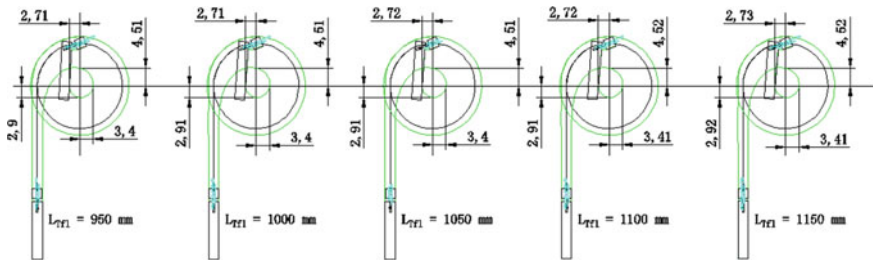


Fig. 8 Simulation results of fifth wheel lead change

-10 mm, -20 mm, -10 mm and the quotients as evaluate indexes are -0.05, -0.05, -0.10, -0.10 corresponding. In conclusion, adjusting the length of fifth wheel lead only seldom affects the passing ability of articulated vehicle.

(3) Tractor Wheel Space

Length of tractor wheel space is 3200 mm, which is changed as previously mentioned only and other parameters remain the same in passing ability simulation. In this condition, length of articulated vehicle don't change. The results of simulation are shown in Fig. 9.

When the length of tractor wheel space changes in the range of 200 mm, the difference of turning clearance circle at different angles of 90°, 180°, 270°, 360° are -100 mm, -130 mm, -150 mm, -100 mm and the quotients as evaluate indexes are -0.50, -0.65, -0.75, -0.50 corresponding. In conclusion, decreasing the length of tractor wheel space reasonably will enhance the passing ability of articulated vehicle effectively.

(4) Semitrailer Wheel Space

Length of semitrailer wheel space is 8100 mm, which is changed as previously mentioned only and other parameters remain the same in passing ability simulation. In this condition, length of articulated vehicle don't change. The results of simulation are shown in Fig. 10.

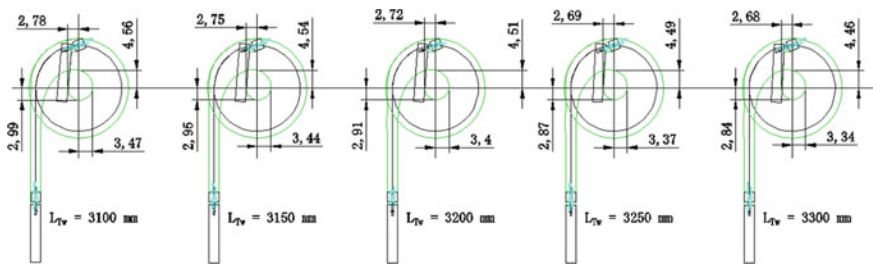


Fig. 9 Simulation results of tractor wheel space change

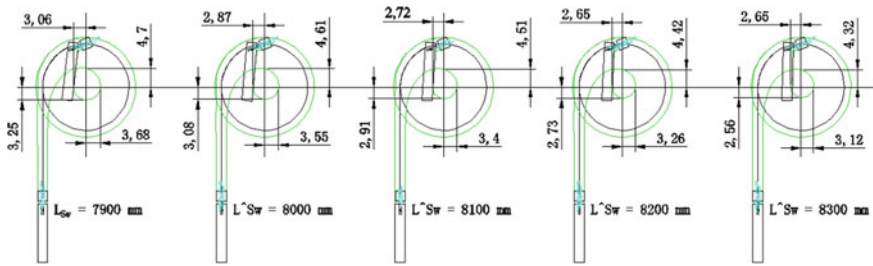


Fig. 10 Simulation results of semitrailer wheel space change

When the length of t semitrailer wheel space changes in the range of 200 mm, the difference of turning clearance circle at different angles of 90°, 180°, 270°, 360° are -380 mm, -560 mm, -690 mm, -410 mm and the quotients as evaluate indexes are -0.95, -1.40, -1.73, -1.03 corresponding. In conclusion, decreasing the length of semitrailer wheel space reasonably will enhance the passing ability of articulated vehicle effectively.

In summary, tractor wheel space and semitrailer wheel space are the main impact parameters in passing ability simulation of articulated vehicle, which is combined of 6 × 4 tractor and van semitrailer. Decreasing the length of tractor wheel space and semitrailer wheel space reasonably will enhance the passing ability of articulated vehicle effectively. In order to improving articulated vehicle passing ability, it is recommended to decrease the length of tractor wheel space 100–150 mm and the length of semitrailer wheel space 150–250 mm.

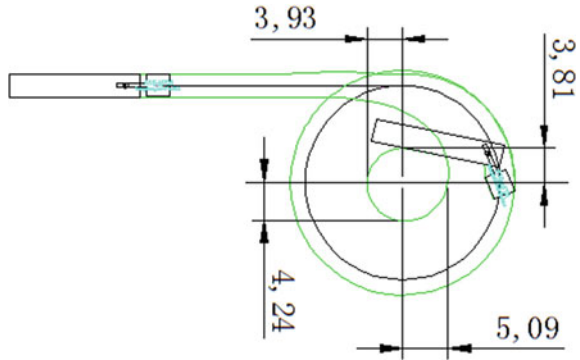
4 Simulation Model Validation

The validity of the simulation model is validated by the test data of ‘3rd Swap Trailer Transport Recommended Vehicles Test’. One 6 × 4 Semitrailer tractor and a container semitrailer and another 6 × 4 Semitrailer tractor and a van semitrailer are selected as compared models, which have the same structure and demission parameters.

4.1 Container Semitrailer

In the test of articulated vehicle combined of tractor and container semitrailer, when vehicle passing the points of 90°, 180°, 270° and 360° on the circle, the corresponding radius are 5700 mm, 4950 mm, 4680 mm and 4570 mm. In corresponding simulation, when vehicle passing 90°, 180°, 270°, 360°, the corresponding radius are 5440 mm, 4720 mm, 4490 mm, 4420 mm, the result is shown in Fig. 11. Corresponding differences are 260 mm, 230 mm, 190 mm and 150 mm, the error between simulation and

Fig. 11 Simulation results of container semitrailer



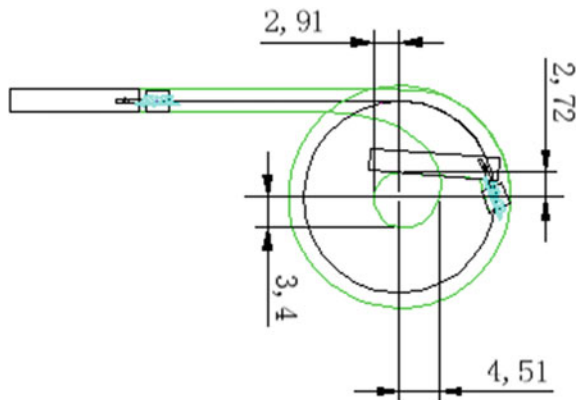
test are 4.56%, 4.65%, 4.06% and 3.28%. Those are all below 5%, thus it proves the validation of simulation model.

4.2 Van Semitrailer

In the test of articulated vehicle combined of tractor and semitrailer, when vehicle passing the points of 90°, 180°, 270° and 360° on the circle, the corresponding radius are 4765 mm, 3680 mm, 3155 mm and 2875 mm. In corresponding simulation, when vehicle passing 90°, 180°, 270°, 360°, the corresponding radius are 4510 mm, 3040 mm, 2910 mm, the result is shown in Fig. 12.

Corresponding differences are 2255 mm, 640 mm, 245 mm and 155 mm, the error between simulation and test are 5.21%, 16.75%, 7.42% and 5.11%. The reason why error between simulation and test in 180° is bigger is that there is a certain different between software model and actual working conditions and measurement error may

Fig. 12 Simulation results of van semitrailer



exist in the road test. Those are all below 5% except the 180° value, thus it proves the validation of simulation model.

5 Conclusions

Passing ability of articulated vehicles is analyzed, which are combined of different 6×4 tractors and different Semitrailers. Articulated vehicles' passing ability simulation is based on requirements of passing ability in GB 1589-2016. Appendix B in this standard gives a method of how to test vehicle turning clearance circle diameter and out value method. The inner radius of each angle is measure value of intermediate shaft of Semitrailer track. Passing ability impact factor of articulated vehicles type and structure are analyzed. Tractor front overhang, distance between center of tractor fifth wheel and tractor first axle, semitrailer first wheel space is selected as the main factors of articulated vehicle passing ability, in order to analyzing passing ability impact factor of articulated vehicles type and structure. Simulation analysis is used to study how tractor front overhang, distance between center of tractor fifth wheel and tractor first axle, and semitrailer first wheel space impact passing ability. The results show that decreasing the length of tractor wheel space and semitrailer wheel space reasonably will enhance the passing ability of articulated vehicle effectively. In order to improving articulated vehicle passing ability, it is recommended to decrease the length of tractor wheel space 100–150 mm and the length of semitrailer wheel space 150–250 mm. The validity of the simulation model is validated by the test data of '3rd Swap Trailer Transport Recommended Vehicles Test'. The results provide a useful reference for improving articulated vehicle passing ability level and promote the development of vehicle industry technology.

Acknowledgements This work is supported by central public welfare research institutes for basic research funds No. 2020-9049 and transportation standard and quota project funds No. 2019-07-071.

References

1. Li J, Che S (2005) Understanding and analysis of GB 1589-2004 'Limits of dimensions, axle load and masses for road vehicles'. *Mach Ind Stand Qual* 4:4–7
2. AutoTURN 9.0. The world's most widely used swept path analysis software [EB/OL]. <http://www.transoftsolutions.com/autoturn>
3. Zheng L, Bai Z, Zhao W (2012) AutoTURN applications in engineering. *Sci Technol Inf* 20
4. Ren Z (2006) Intelligent road design. *Intell Build City Inf* 01
5. GB 1589-2004. Limits of dimensions, axle load and masses for road vehicles
6. Guo Z (2009) Modern combination vehicles design and use. Beijing Institute of Technology Press, Beijing

Roller Opposite Forces Type Automobile Brake Tester Key Technical Parameters' Experimental Research



Yuanpeng Liu, Xiaoping Tong, and Hao Zhang

Abstract Through the study of using opposite forces type desktop method to detect theoretical analysis of automobile braking performance and real vehicle test, proposing about roller opposite forces type automobile brake tester's roller mechanism, roller surface line speed, wheel slip ratio, drive motor power and key technical parameters' indicators and requirements. Research indicates, the diameter of the roller should be controlled in the range of 240–245 mm, the surface linear velocity of the roller surface is 2.2–2.8 km/h; in order to ensure the maximum braking force collected by the brake table, the wheel slip ratio is 25–35%; roller center distance design based on loading capacity, wheel diameter, formed angle, for roller brake table with rated load capacity of 3 t, 10 t and 13 t, the center distance of the roller is 430 mm, 450 mm, 470 mm, and the dimensional error is ± 10 mm. The height difference between the main and auxiliary rollers should not exceed 30 mm, and the dimensional error is ± 2 mm. Considering the influence of loading capacity and roller surface linear velocity on wheel slip ratio, the power reference value of the roller single-side drive motor is given. Through the Calculation of the roller mechanism and use heavy multi-axis vehicle loading brake detection method, for the roller brake table with load detection capability, when installing, the height of the busbar on the auxiliary roller is 40–45 mm from the ground level is well. It is determined that the loading lift height should be no less than 200 mm from the ground level.

Keywords Roller opposite forces type automobile brake tester · Braking force of front axle · Roller mechanism · Formed angle · Roller surface linear velocity · Loading brake detection

Y. Liu · X. Tong · H. Zhang (✉)

Research Institute of Highway Ministry of Transport, Beijing 100088, China

Key Laboratory of Operation Safety Technology on Transport Vehicles Ministry of Transport, Beijing 100088, China

1 Introduction

There are road test method and bench test method to test the braking performance of automobiles. The bench test method is economical, safe, fast and reproducible. It is not restricted by external environmental conditions. It can quantitatively measure the entire braking process of each axle (wheel), which is beneficial to the analysis of the axle braking force value, distribution and the parameters such as the balance state of the wheel braking force and the braking coordination time, and provide the basis for the fault diagnosis of the brake system. Therefore, the bench test method has become the development direction of automobile braking performance detection and diagnosis, and is widely used at home and abroad. At present, most of our country's motor vehicle inspection and testing institutions use roller opposite forces type automobile brake tester (hereinafter referred to as roller brake table) to inspect the braking performance of vehicles, especially suitable for commercial and operational medium and heavy vehicles.

The technical parameters of the roller brake table are the basic conditions to ensure the fairness, accuracy and scientific of the evaluation of the braking performance of the vehicle. This paper determines the key technical parameters through theoretical analysis and actual vehicle test research, and improves the technical level of bench test methods to detect the braking performance of automobiles.

2 Force Analysis

The roller brake table is mainly composed of two sets of left and right wheel braking force test units and a set of indication and control devices. The wheel brake force test unit is composed of a frame platform, a driving device, a roller mechanism, a lifting device, and a force measuring device, etc. Each wheel braking force test unit is provided with a pair of main and auxiliary rollers. When testing the braking performance of the vehicle, the inspected vehicle drives onto the brake tester, and the wheels are placed between the main and auxiliary rollers. It is supported on the frame with a bearing seat, and the axes of the two rollers are kept parallel. The roller is equivalent to an active road surface, used to support the wheels of the vehicle under inspection, and to bear and transmit the braking force. The wheel braking force is obtained by measuring the reaction force of the wheel braking force acting on the force measuring roller. According to the axial load, the roller brake table is generally divided into 3 t level, 10 t level and 13 t level.

In which:

T_{μ}	Brake friction torque;
F_{x1}, F_{x2}	Tangential friction of roller against wheel;
F'_{x1}, F'_{x2}	The tangential reaction force of the wheel to the roller;
T_{f1}, T_{f2}	Wheel rolling resistance torque.

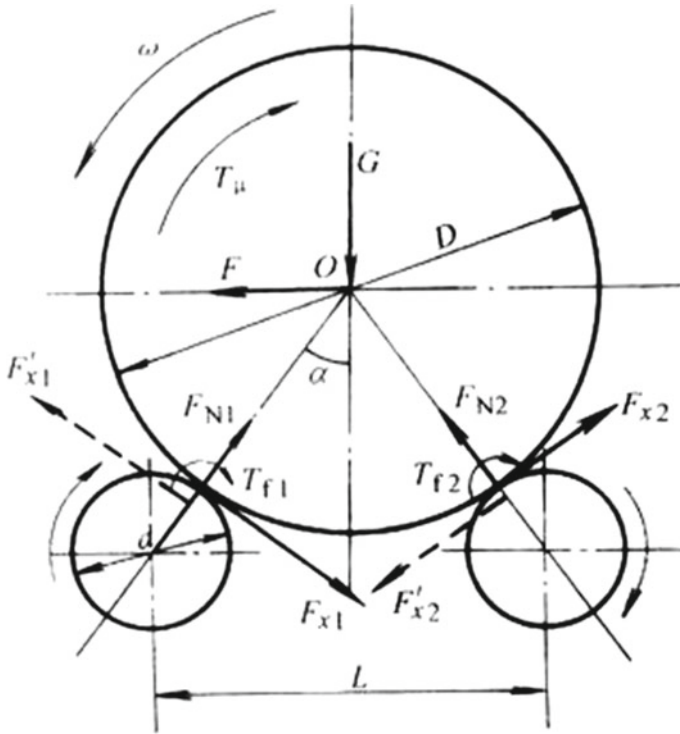


Fig. 1 The forced state of the tested wheel on the roller

Figure 1 shows the forced state of the wheel under test on the roller brake table for brake inspection. According to the principle of mechanical balance, the force of the tested wheel can satisfy the following relationship [1, 2]:

$$F_{N1}(\sin \alpha + \varphi \cos \alpha) - F_{N2}(\sin \alpha - \varphi \cos \alpha) = F \tag{1}$$

$$F_{N1}(\cos \alpha - \varphi \sin \alpha) + F_{N2}(\cos \alpha + \varphi \sin \alpha) = G \tag{2}$$

In Formulas (1) and (2), F_{N1} and F_{N2} are the reaction force of the roller to the wheel; F is the horizontal thrust of the wheel rolling; G is the load on the wheel; α is the installation angle; φ is the sliding adhesion coefficient of the roller. Where the formed angle α is:

$$\alpha = \sin^{-1}(L/(D + d)) \tag{3}$$

In Formula (3), L is the center distance of the roller; D is the diameter of the wheel under test; d is the diameter of the roller.

Solved from the above Formulas (1) and (2):

$$F_{N1} = \frac{F(\cos \alpha + \varphi \sin \alpha) - G(\varphi \cos \alpha - \sin \alpha)}{(\varphi^2 + 1) \sin 2\alpha}$$

$$F_{N2} = \frac{G(\varphi \cos \alpha + \sin \alpha) - F(\cos \alpha - \varphi \sin \alpha)}{(\varphi^2 + 1) \sin 2\alpha}$$

When the wheel brake, the adhesion provided by the roller brake table is:

$$F_\varphi = \varphi(F_{N1} + F_{N2}) = \varphi \frac{G + F\varphi}{(\varphi^2 + 1) \cos \alpha} \quad (4)$$

It is affected by three factors such as the formed angle α , the roller sliding adhesion coefficient φ , and the horizontal thrust F of the wheel rolling (related to the braking performance of the non-detected wheel), When the formed angle α , the roller sliding adhesion coefficient φ , and the horizontal thrust F increase, the adhesion provided by the roller brake table increases accordingly. The formed angle α is related to the diameter D of the inspected wheel, the roller center distance L , and the drum diameter d . When D and d decrease, L increases and the formed angle α increases.

In order to prevent the whole vehicle from slipping when the braking force is detected, and to ensure that the wheel under test cannot be separated from the front roller, that is, $F_{N1} \geq 0$, and $F = 0$, it can be deduced that: $\sin \alpha - \varphi \cos \alpha \geq 0$, that is, $\tan \alpha \geq \varphi$. When the roller sliding adhesion coefficient $\varphi = 0.7$, the corresponding formed angle α should not be greater than 35° .

Through the above force analysis of the inspected wheels, it can be concluded that:

- (1) Properly increasing the formed angle (α) is beneficial for detecting the maximum braking force of the wheels, but it is not as large as possible. Because when the formed angle increases, the relative deformation of the wheel tire increases, the hysteresis loss increases, the additional torque that the roller drives the wheel to rotate increases, which affects the detection accuracy and increases the difficulty of the wheel driving away from the roller. Therefore, the formed angle should be less than 35° .
- (2) The braking force generated by the automobile brake is limited by the ground adhesion coefficient, and its maximum value cannot exceed the ground adhesion [3]. The maximum braking force detected by the roller brake table can only be equal to the adhesion of the wheel under test on the roller brake table (F_φ).
- (3) The size of the roller diameter (D) and the roller center distance (L) has a great influence on the performance of the brake table. Increasing the diameter of the roller is helpful to improve the adhesion between the roller and the wheels, making the detection process closer to the actual braking condition, but the larger the diameter of the roller, the greater the braking force distance (T_μ) generated by the same braking force. Increase the power of the drive motor accordingly. As the diameter of the roller increases, the center distance between the two rollers also needs to be increased accordingly to ensure a proper formed

angle, which will increase the size of the inspection table structure and increase the manufacturing cost.

- (4) When the actual braking force is detected, if the linear velocity (V_t) on the surface of the roller is low, it will affect the rising speed of the braking force of the detected wheel, which will extend the braking coordination time. If it does not match the braking sampling time well, it will affect the detected braking force value. Therefore, when determining the roller surface line speed, the influence on the braking coordination time should be considered.

3 Roller Mechanism

3.1 Roller Diameter

It can be seen from Fig. 1 that when the vehicle under test performs braking detection on the roller brake table, one axle wheel of the vehicle under test is placed on the double-group roller, and the other axle wheel is supported by the ground. Each wheel is in contact with the front and rear rollers at two points (surfaces), which are tangent to the curved surface of the rollers, which is different from the support situation on the ground. At this time, the weight of the wheels is supported by the auxiliary (front) and main (rear) rollers. When detecting the wheel braking force, under the condition that the tested vehicle does not move backwards, the roller sliding adhesion coefficient and the horizontal thrust of the non-tested wheel are fixed, it depends on the formed angle of the tested vehicle on the roller, which can be seen from Formula (3) The formed angle is restricted by the parameters such as the diameter of the roller.

The roller diameter is one of the main structural parameters of the roller brake table. Uniform roller diameter means that the diameter of the roller that determines the formed angle becomes constant, which can solve the problem of data comparability in the detection of different brake table to a certain extent. Compared with the small-diameter roller, the large-diameter roller has a large contact area between the tire and the roller, and the tire deformation is small, which is more conducive to detection. At present, in addition to Japan still producing and using low-speed, small roller brakes, China and European and American countries are dominated by high-speed (roller surface linear speed, relative to Japanese brake table) and large-diameter roller brakes in the market. For the status, the roller diameter of the roller brake table of Europe and the United States is greater than 200 mm.

Years of application practice shows that the diameter of roller of the roller brake table is not less than 240 mm [4] to ensure the contact area between the tire and the roller during the test, and the maximum braking force is collected. At present, domestic equipment manufacturers generally organize design, production and product inspection according to this standard. Taking into account the pipe diameter and production cost of equipment in my country's steel market, the diameter of the roller is determined to be 245 mm \pm 5 mm.

3.2 Roller Center Distance

The roller center distance refers to the distance between the central axle of the main and auxiliary rollers, and is one of the main structural parameters of the roller brake table. It can be seen from Formula (3) that when the roller diameter and the center distance of the roller are determined, the formed angle of the roller brake table is basically determined, and the tire diameter becomes the only variable. Under the condition of a certain roller diameter, the tire diameter is determined according to the rated load quality of the roller brake table, and the limit value of the center distance of the roller is required, which can solve the problem of comparability of the test results of different vehicles in the same vehicle to a certain extent. The formed angle determined according to the above principles is basically consistent with the formed angle range of the foreign roller brake table. For the impact of the center distance of the roller on the braking force detection, we conducted a special test study (see Fig. 2). The test data of passenger cars and heavy vehicles are shown in Tables 1 and 2, respectively.

The above test data shows that with the increase of the roller center distance and the formed angle, the measured braking force increases synchronously, and the fitting curve increases nonlinearly. Obviously, the increase of the roller center distance and the formed angle is beneficial to improve the detection capability of the brake table [5].

In order to meet the needs of vehicle inspection, based on the diameter of the inspected vehicle tires, Formula (3) is used to synthesize the above test data and analysis to determine the roller brake table with rated load masses of 3 t, 10 t, and 13 t. See Table 3 for the roller center distance and error range.



Fig. 2 The experiment of the influence of different roller center distances on the detection of braking force

Table 1 Relationship between center distance of different rollers and shaft braking force (passenger cars)

Model of roller brake table		WZD-13	Roller diameter		245 mm					
Tire specifications		185/65R14	Vehicle type		Buick SGM7150ATA					
Center distance (mm)	Formed angle (°)	Axle braking force (daN)						Axle load average (daN)		
		Once		Twice		Four times				
480	27.31	315	331	322	310	331	313	331	315	642.00
492	28.05	337	320	337	317	337	317	333	322	655.00
500	28.55	346	349	331	315	328	323	328	323	660.75
510	29.18	346	333	328	321	351	336	349	345	677.25
520	29.81	392	342	339	355	350	338	337	309	690.50

Table 2 Relationship between center distance of different rollers and shaft braking force (heavy vehicles)

Model of roller brake table		Roller diameter		245 mm						
Tire specifications		Vehicle type		Heavy semi-trailer towing vehicle Haohan ZZ4255N3246E1						
Center distance (mm)	Formed angle (°)	Axle braking force (daN)						Axle load average (daN)		
		Once	Twice		Three times		Four times			
500	19.05	1506	1299	1515	1292	1480	1268	1524	1276	2790.00
510	19.44	1496	1309	1479	1308	1481	1310	1481	1310	2793.50
520	19.84	1516	1275	1470	1254	1568	1327	1512	1313	2808.75
530	20.24	1512	1274	1513	1332	1482	1260	1527	1309	2802.25
540	20.64	1463	1294	1556	1358	1509	1328	1443	1265	2804.00
590	22.65	1520	1288	1520	1288	1540	1297	1546	1310	2827.25
610	23.46	1584	1362	1545	1301	1527	1292	1584	1362	2889.25
630	24.28	1581	1331	1576	1349	1559	1329	1612	1365	2925.50
645	24.9	1597	1363	1584	1330	1584	1354	1572	1347	2932.75
660	25.52	1630	1435	1694	1509	1648	1456	1642	1410	3106.00

Table 3 The requirements of different bearing qualities on the center distance of the roller

Loading quality (t)	Tire diameter (mm)	Formed angle (°)	Roller center distance (mm)	Error range (mm)
3	≤ 620	$\alpha = \sin^{-1}(L/(D + d))$	430	± 10
10	620–900		450	± 10
13	≥ 900		470	± 10

3.3 Roller Surface Line Speed

Theoretically, since there is no transfer of the center of gravity of the vehicle when the roller brake table is detected, as long as the torque of the brake table drive motor is sufficiently large relative to the friction torque (braking torque) of the vehicle brake under test, regardless of the test speed, The brakes can exert maximum efficiency, and the brake table can measure the maximum braking force, that is, the test speed of the roller brake table has a small effect on the braking force value [6], but it will affect the braking coordination time. The surface line speed of Japanese cogging roller is generally 0.1–0.18 km/h, and the surface speed of European roller is more than 2.0–5.0 km/h.

In order to verify the relationship between the surface speed of the roller surface and the braking force and braking coordination time, we selected passenger cars and heavy trucks to carry out the actual car test. The test data are shown in Tables 4, 5, 6 and 7 respectively. The results showed that:

- (1) The linear velocity of the roller surface (V_t) has little effect on the braking force. The braking force measured by the roller brake table fluctuates within a narrow range with the change of the roller surface line speed, and the braking force value is basically controlled within a relatively stable range.
- (2) The line speed on the surface of the roller is related to the braking coordination time. The higher the line speed on the surface of the roller, the shorter

Table 4 Relationship between roller surface line speed (V_t) and braking force (passenger cars)

Test items		V_t (km/h)						
		0.25	0.5	1.0	1.5	2.0	2.5	3.0
Left rear wheel braking force (daN)	1	179	176	190	189	192	179	197
	2	186	184	185	187	194	173	178
	3	164	172	192	192	195	180	184
	Average value	176	177	189	189	194	177	186
Right rear wheel braking force (daN)	1	153	158	152	156	152	151	142
	2	158	155	154	152	153	143	163
	3	153	152	155	154	157	150	148
	Average value	155	155	154	154	154	148	151

Table 5 Relationship between roller surface line speed (V_t) and braking force (heavy truck)

Test items		V_t (km/h)								
		0.25	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
Left rear wheel braking force (daN)	1	447	438	411	432	444	443	457	462	468
	2	445	443	414	440	449	446	455	464	460
	3	454	454	427	439	443	461	451	461	470
	Average value	449	445	417	437	445	450	454	462	466
Right rear wheel braking force (daN)	1	489	472	421	456	468	476	482	481	479
	2	479	463	428	463	469	466	463	483	494
	3	474	463	434	460	470	471	476	483	506
	Average value	481	466	428	460	469	471	474	482	493

Table 6 Relationship between roller surface line speed (V_t) and braking coordination time (passenger car)

V_t (km/h)		0.25	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5
Braking coordination time (s)	1	0.36	0.18	0.14	0.14	0.16	0.17	0.17	0.16	0.16	0.16
	2	0.37	0.17	0.14	0.16	0.17	0.16	0.17	0.16	0.16	0.17
	3	0.41	0.18	0.16	0.16	0.17	0.17	0.16	0.16	0.15	0.16

Table 7 Relationship between roller surface line speed (V_t) and braking coordination time (heavy truck)

V_t (km/h)		0.25	0.5	1.0	1.5	2.0	2.5	3.0
Braking coordination time (s)	1	0.62	0.38	0.24	0.22	0.20	0.20	0.20
	2	0.64	0.40	0.26	0.22	0.20	0.20	0.20
	3	0.66	0.4	0.26	0.22	0.20	0.20	0.18

the measured coordination time and vice versa. After the line speed reaches 2.0 km/h, the braking coordination time no longer changes significantly with the line speed of the brake roller.

- (3) Increasing the roller surface line speed can more realistically reflect the braking response process, although it is beneficial to the detection of braking table, but it is not the bigger the better. As the roller surface line speed increases, the vibration of the mechanical system of the roller brake table is bound to increased, the system error caused by vibration will also increase.

Considering the braking force, braking coordination time and braking process, it is reasonable to determine the line speed of the roller surface as 2.5 ± 0.3 km/h.

3.4 Height Difference Between Main and Auxiliary Rollers

The height difference between the main and auxiliary rollers can prevent the detected vehicle from moving backward when the vehicle is braked, and at the same time it is convenient to drive away from the brake table after the detected vehicle is detected. We tested the effect of different main and auxiliary roller height differences on wheel braking force. The test data is shown in Table 8. The test results show that the difference in height between the main and auxiliary rollers in the range of 30 mm has little effect on the wheel braking force indication [7]. Therefore, the height difference between the main and auxiliary rollers is determined to be 30 ± 2 mm.

4 Sliding Rate of the Third Roller and Wheels

A third roller with a small diameter, which can rotate and swing up and down, is usually provided between the main and auxiliary rollers of the roller brake table. It is used to detect the surface speed of the wheel under test (see Fig. 3). In the non-detection state, the third roller is kept at the highest position by a spring, and the third roller is equipped with a speed sensor. During detection, the wheels of the vehicle under test are placed on the main and auxiliary rollers, and the third roller is pressed down and kept in reliable contact with it. The control device can measure the linear surface speed of the measured wheel through the third roller speed sensor.

When the wheel under test brakes and the wheel speed drops to near critical locking, the control device calculates the “wheel slip rate %” according to the corresponding electrical signal from the third roller speed sensor [8], and measures when it reaches the specified value, the maximum braking force is measured, and the drive

Table 8 The effect of the height difference between the main and auxiliary rollers on the braking force

Test items		Roller height difference			
		10 mm	30 mm	50 mm	60 mm
		Roller surface line speed			
		2.5 km/h	2.5 km/h	2.5 km/h	2.5 km/h
Left rear wheel braking force (daN)	1	328	305	278	265
	2	331	305	268	257
	3	327	311	269	255
	Average value	329	307	272	259
Right rear wheel braking force (daN)	1	325	325	320	315
	2	328	330	315	311
	3	330	331	315	309
	Average value	328	329	317	312

Fig. 3 The third roller



motor is stopped to prevent the drum from peeling the tire and protecting the drive motor. In addition to the above functions, the third roller can also be used as a safety protection device. When the third rollers of the two brake test units are pressed down at the same time, the drive motor circuit of the roller brake table can be connected.

The wheel slip ratio (δ) is the percentage of the difference between the speed of the measured wheel surface line speed relative to the roller surface line speed and the roller surface line speed, which is expressed as:

$$\delta = \frac{V_t - V_l}{V_t} \times 100\% \quad (5)$$

In Formula (5), δ is the wheel slip rate; V_t is the roller surface line speed; V_l is the wheel surface line speed. It should be noted that when the wheels are braked, because the braking torque is opposite to the driving torque of the driving motor, under the effect of the braking torque, the speed of the driving motor with a small power decreases, resulting in a decrease in the surface speed (V_t) of the roller surface. At this time, there is an error in the calculated slip rate.

The test results of the influence of wheel slip rate (δ) on braking force are shown in Table 9. The test shows that the wheel braking force measured at 30% wheel slip rate is greater than the wheel system measured at 20% wheel slip rate, regardless of the “snap step” or “slow step” due to the effect of the sliding adhesion coefficient on the roller surface power. The wheel slip rate is in the range interval. On the premise of ensuring the safety of the tire, the maximum value is used as much as possible. When the wheel slip rate is 25–35%, the braking force released by the vehicle brake reaches the maximum. The roller brake table can measure the maximum braking force. Therefore, the wheel slip rate when the third roller control drive motor is automatically stopped should be in the range of 25–35%. Increasing the wheel slip rate to stop the range is conducive to detecting the maximum braking force, but excessive wheel slip rate is easy to damage the tire, so the wheel slip rate should not exceed 35%.

Table 9 Effect of wheel slip rate (δ) on braking force (test model: two-axle truck)

Axle order	Wheel slip rate (%)	Braking process	Number of trials	Left braking force (kg)	Right braking force (kg)	Total braking force (kg)	Shaft average braking force (kg)	Axle load (kg)	Shaft braking rate (%)		
Front axle	20%	Snap step	1	1010	879	1889	2086	2841	73.4		
			2	1296	990	2285					
			3	1187	895	2085					
		Slow step	1	829	556	1385	1557				
			2	972	616	1588					
			3	985	715	1700					
	30%	Snap step	1	1344	965	2309	2326	3845	81.8		
			2	1362	1057	2419					
			3	1324	927	2251					
Rear shaft	20%	Snap step	1	1147	836	1983	1959			3845	70.2
			2	1116	799	1915					
			3	1095	885	1980					
		Slow step	1	1440	1311	2751	2700				
			2	1418	1243	2661					
			3	1432	1256	2688					
Rear shaft	20%	Slow step	1	1269	1141	2410	2364	3845	61.5		
			2	1333	1175	2508					

(continued)

Table 9 (continued)

Axle order	Wheel slip rate (%)	Braking process	Number of trials	Left braking force (kg)	Right braking force (kg)	Total braking force (kg)	Shaft average braking force (kg)	Axle load (kg)	Shaft braking rate (%)
			3	1121	1055	2176			
	30%	Snap step	1	1388	1294	2682	2702		70.3
			2	1443	1268	2711			
		3	1448	1267	2715				
		Slow step	1	1282	1068	2650	2429		63.1
	2		1290	1071	2361				
	3		1195	1082	2277				

Remarks: front left wheel weight: 1489 kg, front right wheel weight: 1352 kg, front axle weight: 2841 kg, rear left wheel weight: 2055 kg, rear right wheel weight: 1790 kg, rear axle weight: 3845 kg

5 Drive Unit Motor Power

The roller brake table in China has developed from the original Japanese low-speed cogging roller to the current European-style high-speed sand-sticking roller. Due to the increase in the roller surface line speed and the diameter of the roller, the motor power of the driving device required for the same braking force will also increase. For example, the motor power of the Japanese low-speed roller brake platform is $2 \times 0.7-2 \times 2.2$ kW, and the motor power of the European-style high-speed roller brake platform is $2 \times 3-2 \times 11$ kW.

Tests have shown that if the drive motor power of the roller brake table is not enough, under the action of the wheel braking force, a short-term “lost rotation” will occur, resulting in a small error in the wheel slip rate control, resulting in a large axle load mass. For vehicles, the detection braking force is too small (such as rear engine coach). Due to the short loading time of the braking force detection, the short-term overload capacity and blocked torque of the motor can be fully utilized. When the braking rate of the detection shaft is greater than 60%, the blocked current of the motor will be excessive in a short time. Reasonable determination of the motor rated power can not only ensure the detection capability of the brake table, to ensure error-free detection and judgment, but also avoid the cost and energy waste caused by an excessively large motor. Therefore, the rated power of the unilateral roller drive motor of the roller brake table needs to be specified as [9]:

$$P_d \geq \frac{0.3 \times m_e \times g \times V_t}{1.9 \times 3600} \quad (6)$$

In the formula: P_d is the rated power of the unilateral roller drive motor; m_e is the rated bearing shaft mass of the brake table; g is the acceleration of gravity (take 9.81 m/s^2); V_t is the roller surface line speed.

Statistics show that the locked torque coefficient of motors in my country is usually 2.2. In order not to block the motor rotation, a certain safety factor is required, so it is taken as 1.9. At present, the existing domestic 3 t, 10 t, 13 t rated load quality roller brake table is calculated according to the above Formula (6). When the roller surface line speed $V = 2.5 \text{ km/h}$, then $P_{3t} = 3.22 \text{ kW}$, $P_{10t} = 10.74 \text{ kW}$, $P_{13t} = 13.97 \text{ kW}$, according to the production sequence of three-phase motors in China, the actual roller brake table with rated load quality of 3 t, 10 t and 13 t can be selected to be 5.5 kW, 11 kW and 15 kW respectively.

6 Loading Brake Lift Height

GB 21861-2014 “Motor Vehicle Safety Technical Inspection Items and Methods” stipulates that for trucks with three or more axles, and trailers with dual axle or triple axles installed, For some axles (the last axle and Except for one axle), the loading

Fig. 4 German roller reaction force method braking table loading test



axle braking rate and the loading axle braking unbalance rate should be detected in the loading state, and the lifting loading method is generally used for detection [10], see Fig. 4. Therefore, new requirements are put forward for the installation method and the lifting height of the roller brake table detected by heavy vehicles.

6.1 Installation Height of Roller Brake Table

According to GB 7258-2017 “Technical Conditions for the Operation Safety of Motor Vehicles”, “the front axle of the longitudinal centerline of the motor vehicle (single vehicle) is the front axle”, the first and second steering shafts of a dual-steering vehicle are both front axles, and the no-load axle braking rate should be greater than or equal to 60% [11].

Taking a dual-steering vehicle (three-axle heavy truck) with good braking performance as an example, see Fig. 5. If the roller brake table is installed normally (that is horizontal installation: the busbar on the auxiliary roller and the ground plane are on a horizontal plane), because the first steering shaft is partially overhead by the second steering shaft and the rear axle, that is, the lower bus of the wheel has a

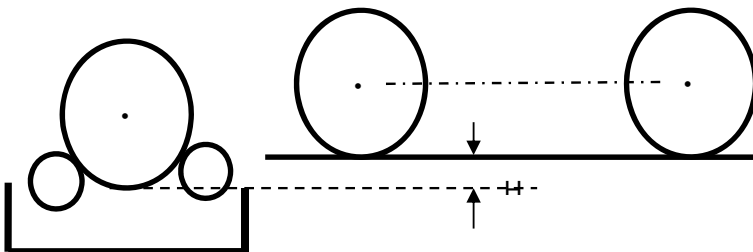


Fig. 5 Schematic diagram of the state of the wheels of the multi-axis car horizontally installing the roller table

sinking amount (H) on the two rollers of the brake platform relative to the ground plane, resulting in a reduction in the mass of the axial load attached to the two rollers, which reduces the maximum braking force and the shaft braking rate detected (relative horizontal weighing), resulting in misjudgment of axle braking table detection. In a similar situation, the semi-trailer towing vehicle and the semi-trailer with parallel axles are detected under the towing of the tractor. The axle load of the inspected axle and the attached axle load of the entire vehicle will also be reduced, which affecting the accuracy of the vehicle braking rate detection [8].

Therefore, it is necessary to standardize the installation method of the heavy-duty roller brake table, that is, the busbar of the auxiliary roller should be higher than the ground plane, and the height of the measured shaft axle should be appropriately increased so that the busbar under the wheel of the measured axle of the multi-axis vehicle slightly higher than the ground plane, which can increase the mass of attached axle load, in line with the actual state of the vehicle, and can effectively solve the problem of brake detection for multi-axis vehicles.

Multi-axis vehicles are mostly heavy vehicles, and the tire diameter is generally 900–1200 mm. Studies have shown that when the vehicle is no-load, when the tire diameter is 1000 mm, the static radius of the steering wheel on the ground is reduced by about 20 mm, and the radius of the driving wheel (and two wheels) is reduced by about 10 mm [12].

Select the roller brake table: the diameter of the roller is 245 mm, and the height difference between the main and auxiliary rollers is 30 mm. The CAD assists the calculation of wheels of different diameters. At different center distances of the rollers, the height (H) of the lower busbar from the wheel to the busbar of the auxiliary roller is shown in Table 10.

(H) generated when the wheel diameter of the heavy-duty multi-axis vehicle is 1000 mm is used to determine the installation height of the busbar on the auxiliary roller relative to the ground plane, and (H) generated when the wheel diameter is 1200 mm is used to verify and control the wheel center of the wheel should not be too high. According to the above calculation, when the roller center distance is 460 mm, the busbar on the auxiliary roller is 40 mm above the ground plane (the error is controlled within 0–5 mm), which can ensure that the wheel axle with a

Table 10 The height from the busbar of wheels with different diameters to the busbar on the auxiliary roller (H)

Wheel diameter (mm)	Main and auxiliary roller height difference: 30 mm		
	Roller center distance (mm)		
	450	460	470
800	37.2	39.6	42.0
900	32.4	34.5	36.7
1000	28.6	30.5	32.4
1100	25.3	27.1	28.9
1200	22.6	24.2	25.8

diameter of 800–1200 mm rises about 10 mm relative to the ground level. The axle of the adjacent wheel rises by about 20 mm (calculated by reducing the static radius of the adjacent wheel by 10 mm). The larger the diameter of the roller, the greater the height of the wheel axle; the greater the difference between the height of the main and auxiliary rollers, the greater the height of the wheel axle.

It should be noted that, according to the regulations of GB 21861-2014, multi-axis and parallel-axle vehicles should be tested for no-load (the busbar on the auxiliary roller is 40 mm above the ground) and load test (the busbar on the secondary roller is 100 mm above the ground). When the installation height of the busbar on the auxiliary roller is increased to the specified value, the loading brake table only needs to be lifted once to complete the loading test. When the busbar on the auxiliary roller is installed flush with the ground, the loading test needs to be lifted twice.

6.2 Loading Brake Lift Height

In order to verify the effect of the load braking lift height on the static axial load, the following tests were carried out:

(1) Double steering axle four axle truck

The test data is shown in Table 11. The test results show that the static axle loads of the first and second axle (steering axle) of a dual-steering four-axle truck increase significantly with the increase in lift, and the third and fourth axle (dual rear axle) are balanced axle. As the lifting height increases, the static axial load does not change much.

(2) Double rear axle three axle truck

The test data is shown in Table 12. The test results show that the static axle load of all axle of the dual rear axle three axle freight car changes relatively little with the increase of lift.

Table 11 Influence of lifting height of four-axle truck (double-turn) on axle load

Horizontal measurement of no-load axial load		First axle	Second axle	Third axle	Four axle
		3241	3232	3309	3161
0 (do not lift)	Axle load (kg)	2425	2010	3002	3019
43 mm	Axle load (kg)	3488	3836	3712	3418
100 mm	Axle load (kg)	4385	5788	3845	3506
200 mm	Axle load (kg)	5152	–	3938	3624
250 mm	Axle load (kg)	5107	/	/	/

Table 12 Influence of lifting height of double rear axle and three axle tractor on axle load

		Lift to 40 mm	Lift to 100 mm
First axle	First time	4883	5173
	Second time	4887	5172
	Third time	4883	5177
Second axle	First time	6073	6466
	Second time	6067	6463
	Third time	6086	6458
Third axle	First time	5825	6073
	Second time	5817	6069
	Third time	5810	6065

(3) Install three-axle trailers (towed by tractors)

The test data curve is shown in Fig. 6. The test results show that the axle load of each axle is increased with the increase of the lifting height. In the early stage of loading and lifting, the axial load ratio (the ratio of loaded axial load to horizontal weighing) increases faster. When lifting more than 200 mm (the height of the busbar on the auxiliary roller from the ground level), the axial load ratio is close to the maximum state, and it shows a slow growth trend.

Based on the above experimental research results are shown as below:

- (1) The roller brake table used for the inspection of multi-axis and parallel-axis vehicles should meet: when the diameter of the roller is 245 mm, the roller center distance is 460 mm, and the height difference between the main and auxiliary rollers is 30 mm, the busbar on the auxiliary roller and the ground level The installation height difference is 40⁰⁺⁵ mm. When the roller center

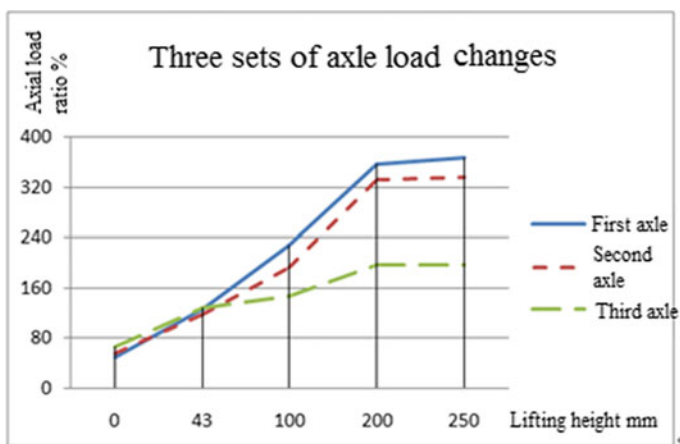


Fig. 6 Test curve of three parallel shafts

distance increases or decreases by 10 mm, the height difference between the busbar on the auxiliary roller and the ground horizontal plane increases or decreases by 2 mm, and when the height difference between the main and auxiliary roller decreases by 10 mm, the height difference between the busbar on the auxiliary roller and the ground horizontal plane corresponds to increase by 4 mm.

- (2) The lifting height of the loading brake table is different, and the measured axial load is also different. The error of the lifting height will affect the brake detection result. When the height difference between the busbar on the installation surface of the roller brake table and the ground horizontal plane is 40 mm, the load detection can be achieved by a lifting height of 200 mm, and under no less than 50% of the rated load mass load, the lifting height error of the loading brake table is not greater than + 5 mm.
- (3) In the lifted state, the attached weight of the inspected wheel increases. For trucks with three or more axles and two-axle or three-axle trailers, the loading brake test is qualified, indicating that the brakes have qualified braking performance. Wheel braking force = wheel attachment weight \times attachment coefficient, the no-load braking test only reduces the weight of the wheels, and the braking table is still qualified. No-load testing is required.
- (4) Under the condition of the maximum lifting height of 100 mm, for the dual-steering four-axle trucks, the first and second steering axles should be tested for loading and braking, and the third and fourth axles (dual rear axles) for lifting and loading braking detection is of little significance. Similarly, for a dual-steer three-axle truck, only the first and second steering axles need to be tested for loading and braking; for all axles of dual-axle three-axle trucks, the lifting load braking test is of little significance; for vehicles with parallel axles, the lifting load test should be implemented for all parallel axles.
- (5) When the tested axle has a maximum lifting height of 100 mm, the reduction in the attached weight of its adjacent axle is not enough to make it large enough to adapt to the load detection of the shaft with air suspension. Therefore, the height of the loading brake detection lift should not be less than 200 mm, so that the adjacent shaft of the detected shaft is close to or completely off the ground, thereby increasing the static axial load of the detected shaft, and solving the problem of loading detection of vehicles with air suspension. At this time, all axles of multi-axle and parallel-axle vehicles should be loaded and tested.

7 Conclusion

Through the theoretical analysis of automobile braking table testing and the experimental study of testing the braking table of the vehicle with the bench method, comprehensively considering the types and uses of in-use vehicles across the country, and the different bearing capacity of the roller brake table, the study proposed key technical parameters of roller opposite forces type automobile brake tester: the roller

Table 13 Key technical parameters of roller brake table

Technical parameters	Rated load quality (m) t		
	3	10	13
Roller diameter (mm)	245 ± 5		
Roller center distance (mm)	430 ± 10	450 ± 10	470 ± 10
Roller surface line speed (km/h)	2.5 ± 0.3		
Height difference between the main and auxiliary rollers (mm)	30 ± 2		
Wheel slip rate (%)	25–35		
Power of the drive motor (kW)	2 × 5.5	2 × 11	2 × 15
Lifting height of the loaded roller brake table (mm)	Roller brake table installation: the height of the busbar on the auxiliary roller from the ground level is 40 ₀ ⁺⁵ mm		
	Each lift height: 200 ₀ ⁺⁵ mm		

diameter, the roller center distance, the roller surface speed, the height difference between the main and auxiliary rollers, the wheel slip rate, the power of the drive motor, and the lifting height of the loaded roller brake table (see in Table 13).

The results of this test can provide technical support for improving national and industry automotive performance testing standards, unifying and standardizing the quality of automotive brake table products, and can also effectively guide the product design, production, use and product inspection of automotive roller brake table.

Acknowledgements This work is supported by transportation standard and quota project funds No. 2018-05-046 and No. 2019-10-073.

References

1. Yu Z (1990) Automotive theory. Mechanical Industry Press, Beijing
2. Cui J, Zhou Z (1995) Comprehensive car performance testing. Shanghai Scientific and Technical Literature Publishing House, Shanghai
3. Xu H (2014) Automotive application engineering. People's Communications Press, Beijing
4. Roller opposite forces type automobile brake tester: GB/T 13564-2005. China Standard Press, Beijing
5. Zhao Z (2008) Research on the variable roller wheelbase opposite forces type brake tester. Jilin University, Changchun
6. Tong X (2006) Experimental research on detecting the influence parameters of automobile braking performance based on opposite forces type brake tester. Automobile Maintenance Repair 2006(08):35–37
7. Miao Z (2005) Post training materials for automotive testing personnel. People's Communications Press, Beijing
8. Tong X (2016) Comprehensive performance inspection and technical grade assessment of road transportation vehicles. People's Communications Press Co., Ltd., Beijing

9. Road transport vehicle comprehensive energy requirements and inspection methods: GB 18565-2016. China Standard Press, Beijing
10. Motor vehicle safety technical inspection items and methods: GB 21861-2014. China Standard Press, Beijing
11. Motor vehicle operation safety technical conditions: GB 7258-2017. China Standard Press, Beijing
12. Research Institute of Highway Ministry of Transport (2010) Research on the maintenance evaluation system of technical status of operating vehicles. Research Institute of Highway Ministry of Transport, Beijing

Trends Analysis of Domestic Waterway Traffic Safety Production Based on Regional Characteristics in China



Guobo Wang, Guanquan Chu, and Chao Han

Abstract In order to improve the level of waterway traffic safety production in China, based on the significant differences in waterway traffic safety production in different regions, analogy and trend analysis methods are used to analyze the safety production situation of water transportation in different regions from multiple perspectives such as shipping, ports, ships, and accidents. The results show that the waterway freight volume in different areas maintains a relatively stable growth, and the long-distance transportation of bulk cargo is the main development trend. The throughput growth of inland river ports slowed down, and the throughput of coastal ports continued to increase. The average deadweight of transport ships in inland water shows a significant linear growth trend, while the passenger capacity of ships shows a fluctuating downward trend. The distresses are mainly concentrated in coastal waters, and the characteristics of the distresses in different areas are different. The number of distress in coastal waters is mainly related to the number of severe weather, while the number of distress in Inland River has the highest correlation coefficient with the number of ships.

Keywords Waterway traffic · Safety production · Regional characteristics · Correlation coefficient

1 Introduction

Transportation is a basic, leading, strategic industry and service industry for economic and social development. With the continuous and rapid development of the national economy and the increasing demand for people's transportation, although the total number of safety production accidents in the transportation industry has shown a continuous and steady decline, the total number of accidents is still large and the situation of safety production remains severe [1, 2].

G. Wang · G. Chu · C. Han (✉)
China Waterborne Transport Research Institute, Beijing 100088, China
e-mail: hanchao@wti.ac.cn

Based on the literature research, it can be found that the existing research mainly focuses on waterway traffic safety and development trends, which mainly use accidents and distresses as measurement standards. While the waterway transport production involves the entire waterway transportation process, and the safety analysis of the whole process is more practical [3, 4]. In particular, the domestic waterway transportation system is intricate, and the status quo of waterway traffic safety production in different regions is significantly different. Therefore, it is necessary to carry out analysis of the waterway traffic safety production situation in different regions. At present, waterway transportation is mainly concentrated in three aspects: inland water transportation, coastal transportation and ocean transportation. Due to the different water and meteorological conditions in the area where these three transportation modes are located, there are also large differences in their impact on the safety production of waterway transportation. Therefore, regional research on safety production of waterway transportation is carried out to provide scientific guidance for improving the level of safety production in different regions of China.

2 Shipping Development

In 2019, the navigable mileage of inland channels reached 127,300 km, an increase of 0.16% compared with last year. From the perspective of the changes in the navigation mileage of inland channels in the past ten years, the total navigable mileage of inland channels has been slowly increasing every year, and has become saturated after 2015, stabilizing at around 127,000 km. The mileage of graded channels is also increasing, exceeding the mileage of non-grade channels for the first time in 2010. The mileage of non-grade channels is showing a trend of declining. This is related to the strengthening of the Yangtze River waterway regulation by the transportation authorities, and the navigation safety of the Yangtze River waterway has been effectively improved [5, 6] (Fig. 1).

In 2019, the freight volume of inland water transportation, coastal transportation and ocean transportation was 3.913 billion tons, 2.727 billion tons, and 832 million tons respectively. Inland water transportation accounted for 52.37% of the freight volume, which is the main method of waterway cargo transportation. From the perspective of freight turnover, due to the long haul distance of ocean transportation, the freight turnover is relatively large. The freight turnover of inland water transportation is relatively small because of the short haul distance. As shown in Fig. 2, the gap in freight turnover between ocean transportation and inland water transportation is getting smaller. This is mainly due to the upgrading of ships, the increase in the mileage of inland waterways and the continuous expansion of inland waterway routes, resulting in average haul distance and freight volume continue to increase. It can be seen from Fig. 3 that the freight volume of inland water transportation, coastal transportation and ocean transportation has maintained a relatively stable growth trend in the past ten years. The average growth rate of freight volume of inland water transportation and coastal transportation is significantly higher than

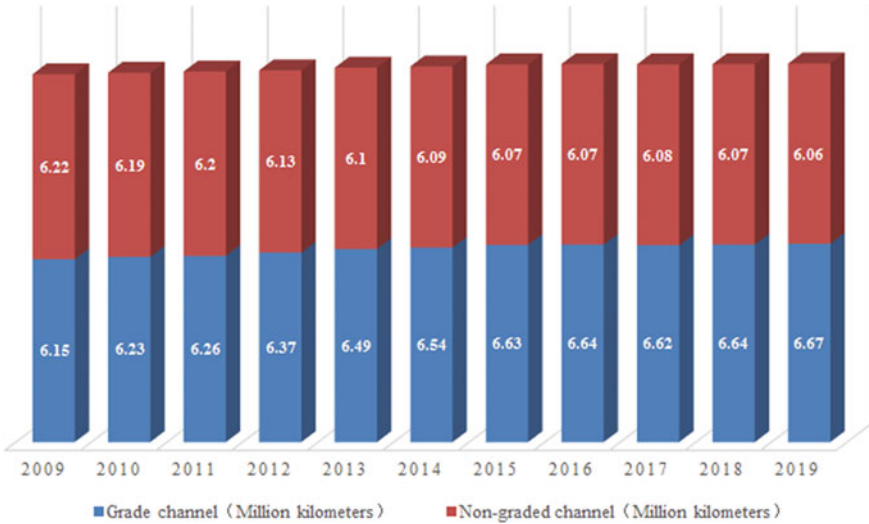


Fig. 1 The navigable mileage of inland channels from 2009 to 2019

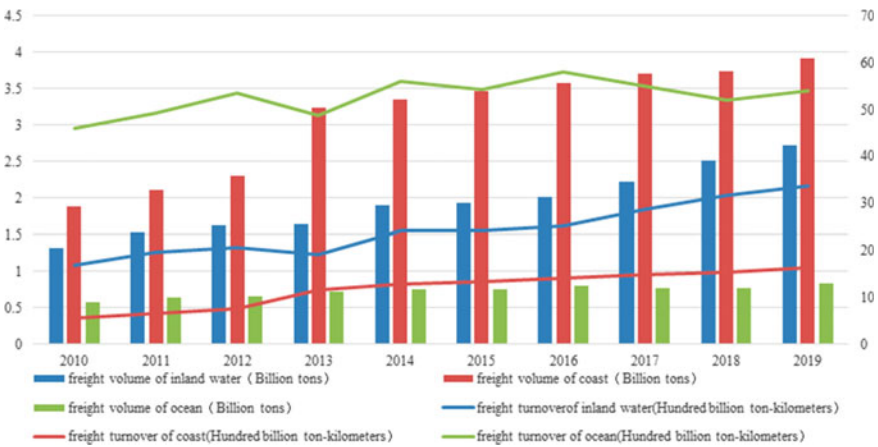


Fig. 2 The freight volume and turnover of waterway transportation in different regions from 2010 to 2019

that of ocean transportation. The average haul distance of coastal transportation is also increasing, indicating that long-distance transportation of bulk cargo is the main trend in the development of waterway transportation in the future, which is also the changing trend of waterway transportation safety.

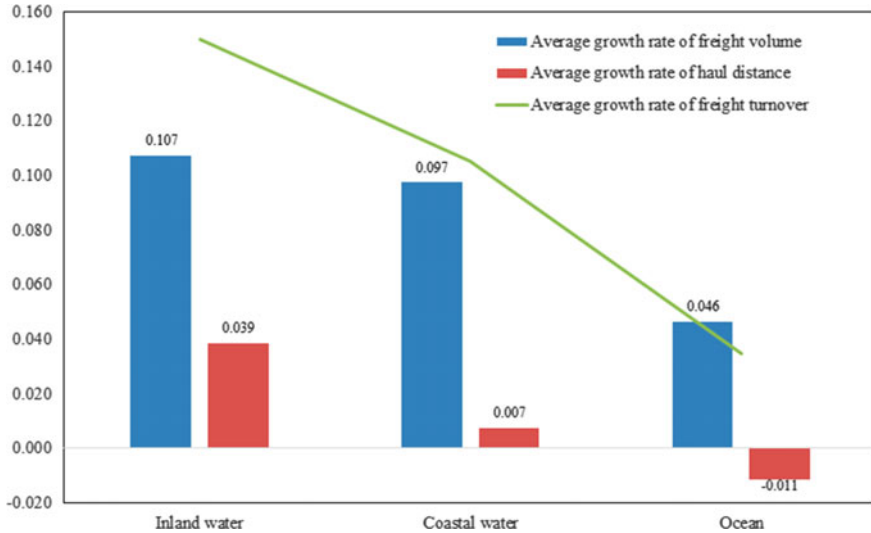


Fig. 3 The average growth rate of freight volume, average haul distance, and freight turnover from 2010 to 2019

3 Port Development

At the end of 2019, there were 22,893 production berths nationwide, a decrease of 1026 from the previous year. Among them, there were 5562 berths in coastal ports, a decrease of 172; and 17,331 berths in inland ports, a decrease of 854. There were 2520 berths over 10,000 tons, an increase of 76 over the previous year. Among them, there were 2076 berths at 10,000-ton level and above in coastal ports, an increase of 69; and 444 berths at inland ports with 10,000-ton level and above, an increase of 7. Judging from the changes in the number of berths in the past ten years (Fig. 4), the number of berths has first increased and then decreased. The number of berths reached the highest in 2011, and the number of inland berths was nearly five times that of coastal berths. After 2011, the total number of berths began to decrease, especially in the past three years, with an average decrease of 11.26%. As of the end of 2019, the number of inland berths has become about three times the number of coastal berths. The change in the number of inland berths is basically the same as the total number of berths, while the number of coastal berths has been on a steady growth trend, with only a slight decrease in 2019.

It can be seen from the changes in the number of 10,000-ton berths in the past decade (Fig. 5) that the number of 10,000-ton berths has been increasing, and this growth trend has a very significant linear characteristic. Similarly, the increase in the number of 10,000-ton berths in coastal waters and inland rivers both showed strong linear characteristics. The fits were 0.99 and 0.92, and the slopes were 87.345 and 14.182. It can be seen that both the number of 10,000-ton berths and the growth rate in coastal waters are higher than inland rivers, and their scale effect is much higher

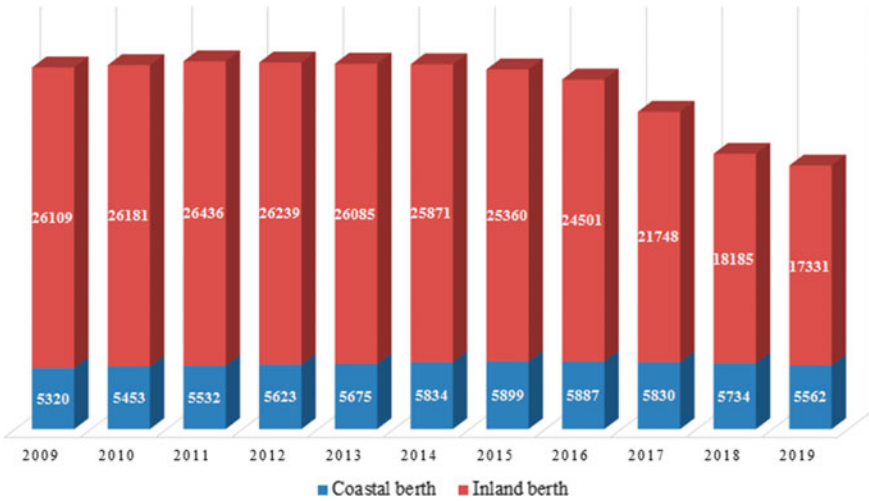


Fig. 4 Number of berths from 2009 to 2019

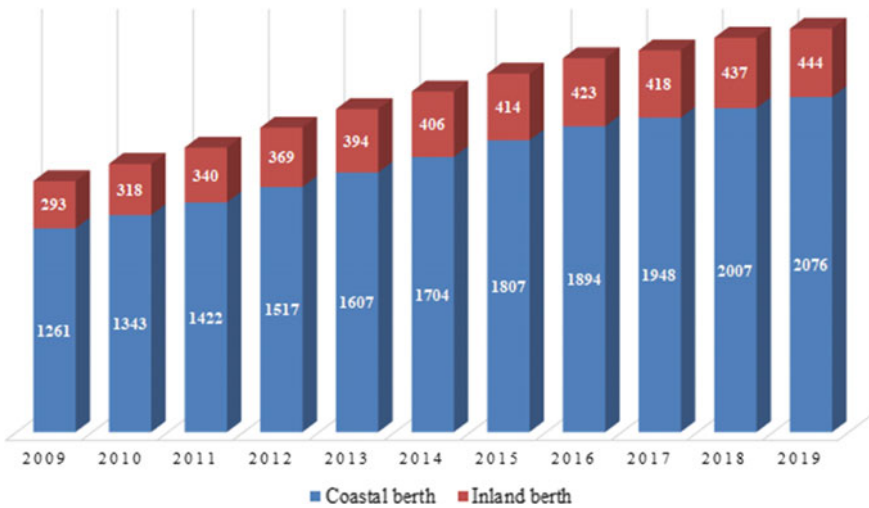


Fig. 5 Number of 10,000-ton berths from 2009 to 2019

than inland rivers. The average growth rate of the number of 10,000-ton berths in the past ten years is 5.63%, 5.44% for inland rivers, and 5.68% for coastal. Combined with the linear characteristics of the fitting, it can be seen that although the average growth rate of the number of 10,000-ton berths in inland rivers and coastal waters is the same, the growth trend of the number of 10,000-ton berths in coastal waters is getting better and better, while the growth rate of the number of 10,000-ton berths in inland waters tends to be flat. This is mainly due to the limited navigation and

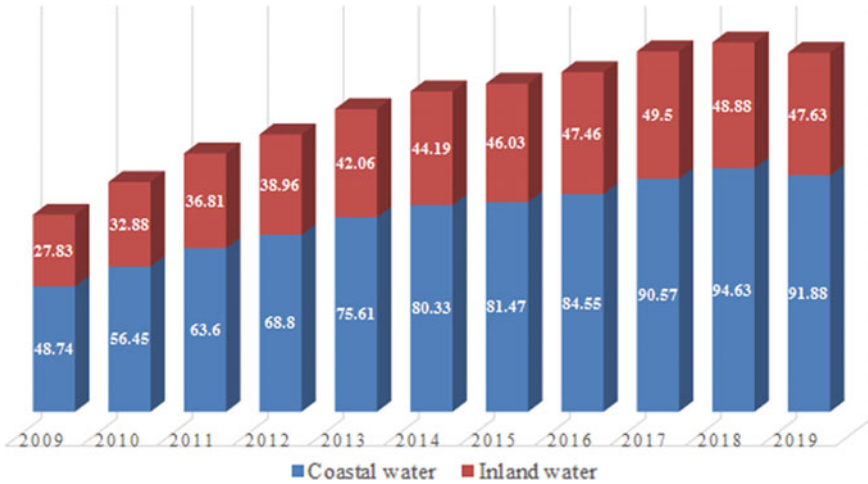


Fig. 6 Port cargo throughput from 2008 to 2018

terminal conditions in the inland waters, as well as the tightening of approvals for wharf renovation, which has slowed down the growth rate of 10,000-ton berths in inland waters.

In 2019, all ports across the country completed a cargo throughput of 13.951 billion tons, a decrease of 2.9% from last year. Among them, coastal ports completed 9.188 billion tons, a decrease of 2.9%; inland ports completed 4.763 billion tons, a decrease of 2.6%. As shown in Fig. 6, the port cargo throughput has maintained a relatively stable and continuous growth in the past ten years, and its growth trend has a strong linear characteristic, which is more consistent with the growth trend of the number of berths. The cargo throughput of coastal and inland river ports also showed a linear growth trend, and the growth trend of coastal ports was better than that of inland river ports, which was basically consistent with the growth trend between the number of coastal and inland river berths. The average growth rate of port throughput in the past ten years is 7.49%, 6.91% for inland rivers, and 7.82% for coastal areas. Combining the linear characteristics of the fitting, it can be seen that the port throughput of inland waters has increased significantly in the past few years.

4 Transport Ship

At the end of 2019, there were a total of 131,600 transport vessels nationwide, a decrease of 3.9% over the previous year; a net load of 256.8497 million tons, an increase of 2.3%; a capacity of 885,800 passengers, a decrease of 8.0%. The changes in the number of ships in the past ten years can be seen from Fig. 7. The total number of ships has basically shown a downward trend. In particular, the downward



Fig. 7 Number of water transport vessels from 2008 to 2018

trend in the past five years has been particularly pronounced. The average decline rate has reached 5.5%, and the number of ships that have decreased is mainly inland waterway vessels. This is mainly because the state has strengthened the management of old ships in order to optimize the structure of waterway transportation capacity, improve the technical level of ships, and ensure the safety of waterway transportation. At the same time, the influence of other transportation modes such as roads and railways continues to expand, which were further reducing the living space of water transport ships, resulting in a continuous decrease in the number of inland watercraft. Compared with inland water transport ships, the number of coastal transport ships has little change. After experiencing a small increase, it began to decrease slowly in 2015 and stabilized. Due to the high construction requirements of maritime transport ships, specialized ships are the main focus. Meanwhile, the replacement of ships is fast, and the overall number is small, so the overall change is not obvious. Ocean-going ships are limited by cargo volume and transportation demand, the number of ships is small, and basically maintaining a relatively stable number.

It can be seen from the change trend of the ship's net deadweight and average deadweight in the past ten years (Fig. 8) that the ship's net deadweight has shown a trend of increasing first and then decreasing, while the average deadweight of ships has always shown a slow growth trend, and has a relatively significant linear characteristic, with a correlation coefficient of 0.97. It shows that the trend of large-scale ships is constantly expanding. With the number of ships shrinking, the net weight of ships can still maintain a relatively gentle rate of reduction. The net deadweight of inland and ocean ships has basically the same trend as the net weight of ships nationwide. However, compared with the trend of the average net deadweight of

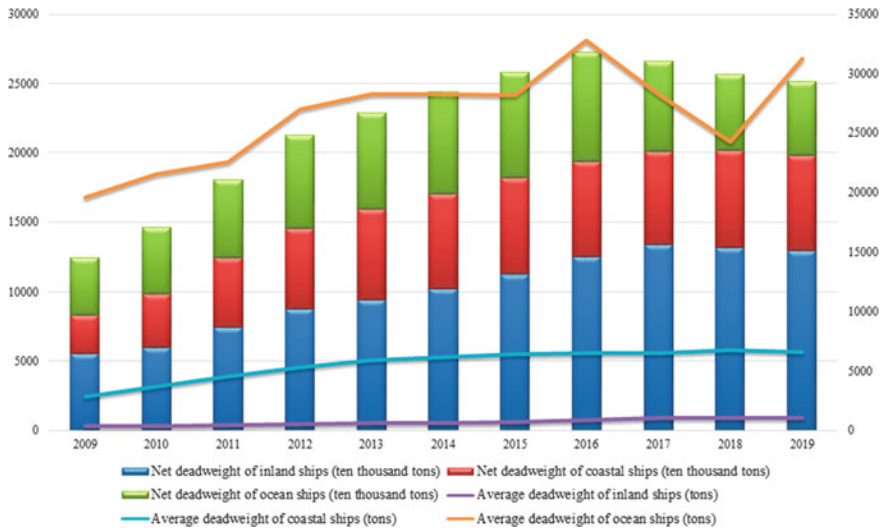


Fig. 8 The net deadweight and average deadweight of ships from 2008 to 2018

ocean-going ships, the average deadweight of inland water transport ships showed a significant linear growth trend, with a correlation coefficient of 0.99. It shows that the large-scale effect of inland watercraft is relatively significant, and the increase in large-scale ships has better compensated for the impact of the decrease in the total number of ships. As the number of ships and the demand for freight volume decreases, the average deadweight of ocean-going ships has continued to decline, and it has leveled off in 2017.

It can be seen from the trend of the passenger capacity of ships in the past ten years (Fig. 9) that the passenger capacity of ships has shown a fluctuating downward trend, but the decline is relatively low, the passenger capacity of ships nationwide is relatively stable. The passenger capacity of inland waterways has basically shown a declining trend, and the downward trend has been obvious in the last three years. This is mainly due to the substitution effect of railways, highways and other transportation modes, and the decline in water passenger traffic has led to the cancellation of some waterway routes. Compared with inland rivers, the passenger capacity of coastal ships has shown a steady upward trend. The main reason is the continuous improvement of ship technology and the development of coastal tourism routes. The number of tourists choosing cruise tourism has gradually increased, and the demand for coastal passenger transportation has gradually increased. Due to the natural disadvantages of long distances and long cycles, ocean-going routes have limited demand for ocean-going passenger transport, and their passenger capacity is relatively stable.

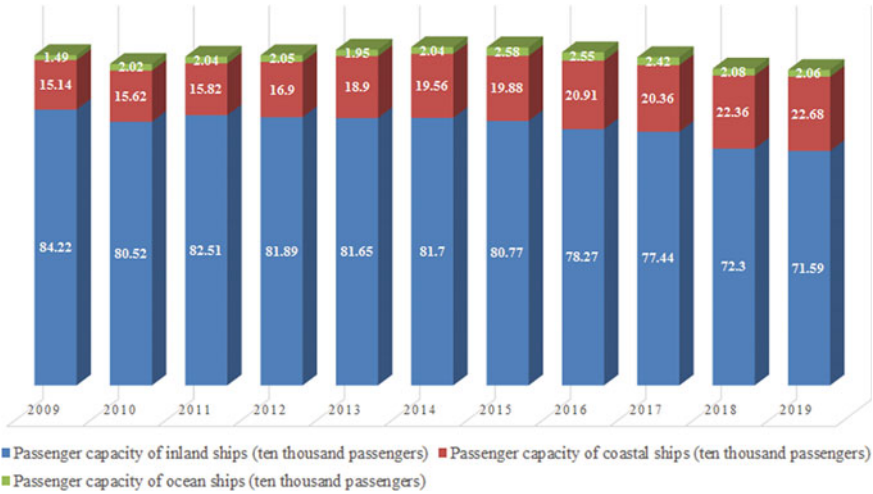


Fig. 9 Passenger capacity of ship from 2008 to 2018

5 Traffic Safety on Waterway

The domestic navigable waters mainly include coastal areas, main stream of inland rivers, and tributary of inland rivers. Therefore, the safety of water traffic is analyzed from the above three areas. In 2019, there were 2,240 distresses in navigable waters across the country, including 1,564 distresses in coastal waters, 230 distresses in main stream of inland rivers, and 446 distresses in tributary reservoirs. As shown in Fig. 10, the distresses are mainly concentrated in coastal waters, accounting for 73% of the total number of distress in the past five years. This is mainly due to the large number of navigable risk factors in coastal waters and harsh weather conditions. In addition, the number of distress in coastal waters is on a fluctuating upward trend, and it is necessary to pay sufficient attention to the navigation safety of coastal waters, especially the safety of ships and personnel search and rescue under severe weather conditions need to be further improved. The number of distresses in the main stream of inland rivers has shown a relatively steady decline. It is possible that after 2013, the inland waters began to carry out water transportation construction and improvement work [5]. The safety situation of inland waters has gradually improved, which has led to a gradual reduction in the number of distresses. Although there are few navigable ships in the tributary reservoirs of inland rivers, the relatively poor navigable conditions and the relatively large proportion of non-standard ships have led to the occurrence of distresses [6]. The number of distresses is small, but it is still necessary to strengthen water safety management in waters with weak navigable conditions.

In 2019, a total of 528 people died in navigable waters across the country due to distresses, including 446 people in coastal waters, 45 people in inland waters, and 37 people in tributary reservoirs. As shown in Fig. 11, coastal waters accounted for the

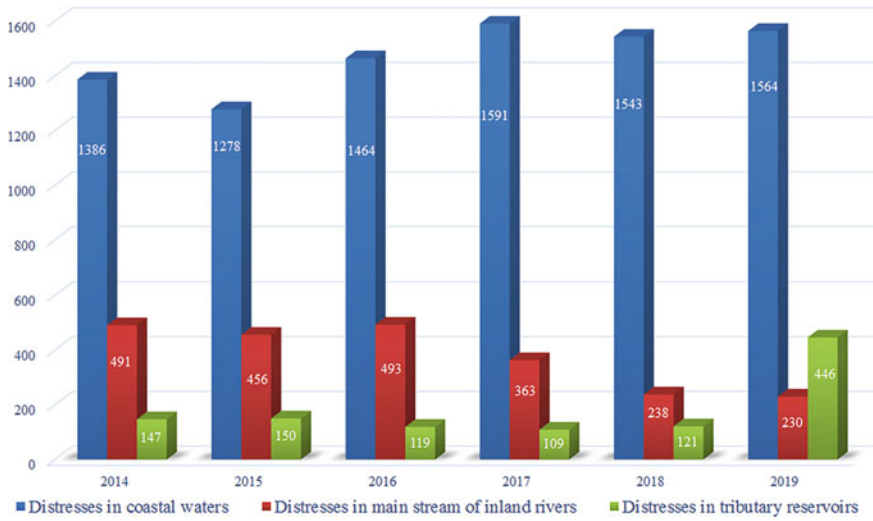


Fig. 10 Distresses in navigable waters from 2014 to 2019

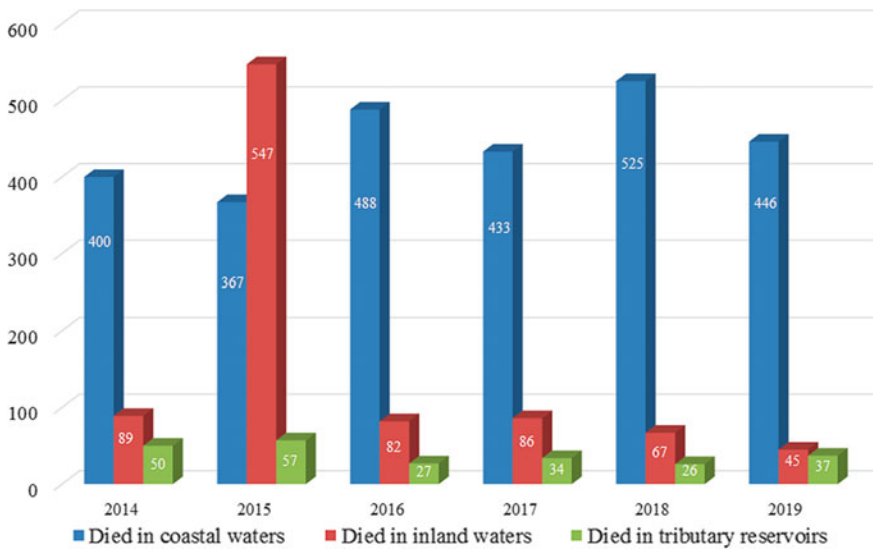


Fig. 11 Number of deaths in distresses from 2014 to 2019

largest number of deaths, accounting for 67.5% of the total deaths in navigable waters in the past five years. The trends in the number of deaths and the number of distresses are basically the same. The main reason is that the environmental conditions of coastal waters are poor. Once a dangerous situation occurs, rescue is more difficult, and the probability of survival of people in distress is small, which leads to an increase

in the number of deaths. The death toll of the inland waters as a whole is relatively small, basically showing a slow decrease trend. Except for the large number of deaths caused by the fall of the “Oriental Star” in 2015 [7]. It shows that the overall safety of the inland waters is relatively good, but since the passenger volume in the inland waters accounts for a relatively large amount, sufficient attention should be paid to the management of passenger ships. Attention also should pay to the safety of passenger ships in coastal waters attention. Due to the rapid development of cruises and the difficulty of search and rescue in coastal waters, once a cruise ship overturns or other water traffic accidents occur, the consequences will be disastrous. The relevant departments need to strengthen the management of passenger ships and actively guide a healthy and stable development of the passenger transportation industry.

By comparing the linear correlation among shipping development, port development, transport ships and distresses in waterway, it is found that the main factors affecting the safety of water traffic in different regions are significantly different (Fig. 12). In coastal waters, there is little correlation between the volume of freight, passenger traffic, and the number of ships and the number of distress. Among all the factors, the correlation coefficient between the number of severe weather and the number of distress is the highest, indicating that ship distress is mainly related to severe weather in coastal waters, which is related to the high difficulty of rescue and distress finding in coastal waters. In the inland waters, the linear correlation between the number of ships and the number of distress is strong, reaching 0.94, which shows that water traffic accidents are mainly caused by poor water traffic conditions in the inland waters. Channel resources in inland waters are limited, and the reduction in

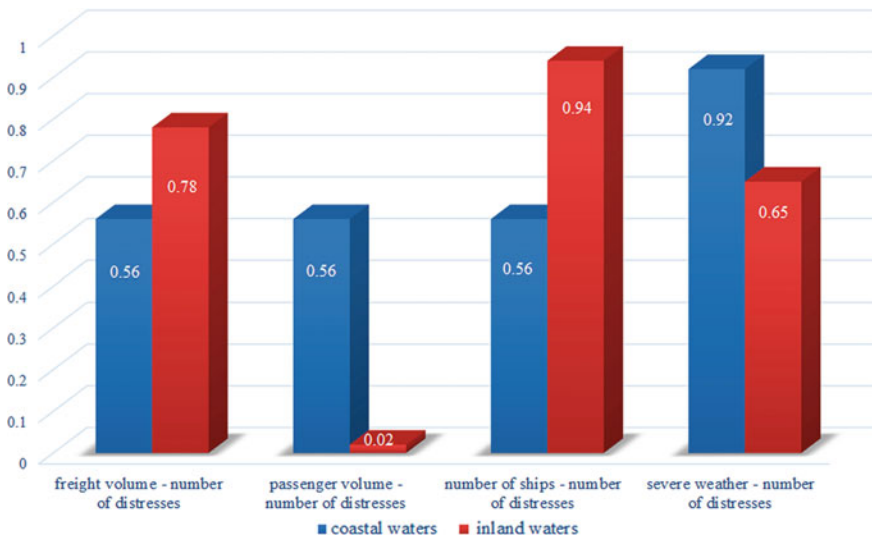


Fig. 12 The linear correlation between different factors in coastal and inland waters

the number of ships also reduces the density of ships in the unit navigable water area, thereby reducing the probability of accidents such as ship collisions.

6 Conclusion

There are significant differences in the influencing factors between coastal and inland waterway traffic safety. In order to improve the level of waterway traffic safety in different regions, it is necessary to suit the remedy to the case based on the characteristics of different regions. For coastal waters, it is necessary to pay more attention to the safety of waterway traffic under severe weather conditions, especially to continuously improve the early warning and forecast of severe weather and the level of emergency search and rescue capabilities at sea. For inland waters, more attention should be paid to the navigation safety of ships in the waterways, safety supervision should be strengthened in waters with a large number of ships, and investment in inland water transport construction should be maintained or increased to improve the hardware conditions and management level of inland waters.

References

1. Xu P, Huang H, Dong N (2018) The modifiable areal unit problem in traffic safety: Basic issue, potential solutions and future research. *J Traffic Transp Eng* 5(01):73–82
2. Grinyak VM, Devyatisilnyi AS, Ivanenko YS (2019) Traffic safety assessment in water area. *IOP Conf Ser Earth Environ Sci* 272:022017
3. Jun T, Xi C, Xiaojun C et al (2018) Research on the key technology optimization of drainage system in production testing project of Yuanba gas field. *Sino-Global Energy*
4. Adams AJ, Johnson MA, Ryan KA et al (2019) Safe transportation in-Spica following surgical treatment of infantile DDH: solutions and threats. *J Pediatric Orthopaedics*
5. Lin X, Wu H, Jiangsu MSA (2018) Consideration on the new safety theory of water transport on the base of “12345” scientific safety supervision system of Jiangsu MSA. *China Mar Saf*
6. Liu, Lijun, Wen et al (2019) Extreme weather impacts on inland waterways transport of Yangtze River. *Atmosphere*
7. Wang X, Zhang B, Zhao X et al (2020) Exploring the underlying causes of Chinese Eastern Star, Korean Sewol, and Thai Phoenix Ferry accidents by employing the HFACS-MA. *Int J Environ Res Public Health* 17(11):4114

Study on Route Optimization of Seasonal Hot Spots



Haixia Zhao, Ting Tong, and Jiangping You

Abstract With the adjustment of holiday policy and the rapid growth of consumers' demand for tourism, severe overcrowding in hot scenic spots has become one of the main bottlenecks that perplex the development of many scenic spots in China. Based on the relevant literature and relevant data obtained from practical investigation, the factors that affect the path planning of hotspots are reasonably weighted by entropy method, and short-term prediction method is used to predict the passenger flow. Based on the improved A* path search algorithm and careful consideration in the various costs of different scenic spots in the scenic area, we can determine which scenic spot should come next, and finally get the optimal path. When tourists arrive at different scenic spots, they will make further progress according to the overall situation of the scenic area and update the optimal path. This paper takes Huangguoshu Tianxing Bridge scenic spot as an example to verify the improved A* path search algorithm. When taking Tianxing Bridge as the starting point, the optimal path is obtained by predicting the passenger flow of each scenic spot in the scenic area and comprehensively evaluating the influence of various factors, taking into account the total cost of each scenic spot. This paper considers the single way for tourists to visit—walking. In fact, for the problems of scattered scenic spots and considerable distance between them, sightseeing cars and shared bicycles would be arranged in the scenic spots. But by changing the value of traffic speed, the arrival time of each scenic spot can be calculated. The result has wide applicability, and the optimal path planning method has certain flexibility, which can be widely used in the route planning of other scenic spots.

Keyword Overcrowding in hot scenic spots · Path planning · Entropy method · Short-term prediction method · Improved A* search algorithm

H. Zhao · T. Tong (✉) · J. You
Southwest Jiaotong University Hope College, Rail Transit Institute, Chengdu, China
e-mail: tongtingTTa@163.com

H. Zhao
e-mail: zhaohaixiahit@126.com

1 Problem Demand

Tourism congestion has become one of the bottlenecks in the development of tourism due to the adjustment of holiday policy and the increase of national consumption [1]. Sceneries are hot season tourists increased dramatically, internal distribution is not balanced, the scenic spots in between is not clear and failure to make timely adjust, and scenic emergency sightseeing car arrangements exist such problems as unreasonable, tourist route scheduling error, unexpected events occur, can make the scenic spot in resource utilization, which in turn results in uneven distribution of tourists and produce structural congestion phenomenon. The model of this project can maximize the comprehensive utilization efficiency of scenic resources and the efficiency of natural ecological protection, finally evacuate the crowded people, reasonably arrange tourists to visit the scenic area, and realize the sustainable development of the scenic area.

2 Data Uses

2.1 Data Survey

2.1.1 Distribution of Huangguoshu Scenic Spots

Huangguoshu Scenic spot covers an area of 115 km² and is rated as the largest waterfall group in the world by the Great World Guinness Headquarters. It is included in Guinness Records of the world. It is composed of Huangguoshu waterfall, Doupotang Waterfall, Dashuyan Waterfall, Dishuitan Waterfall, Luoshitan Waterfall, Yinlianzhuitan Waterfall and other Waterfalls. Each waterfall is surrounded by a variety of scenic spots, large and small. The distribution map of Huangguoshu Scenic spots is shown in Fig. 1.

2.1.2 Actual Distance Between Attractions

1. The distance between the scenic spots from the Huangguoshu is shown in Table 1.
2. The distance between the various attractions from the Tianxing Bridge is shown in Table 2.



Fig. 1 Map of various scenic spots of Huangguoshu

2.1.3 Holiday Flow Statistics

2013: 2013-10-03 began the Golden Week passenger flow exceeded 200,000;

2015: Capacity control after 30,000 people on a single day, Huangguoshu stopped selling tickets for the day. In Huangguoshu Scenic Area, the maximum carrying capacity of tourists in the peak season is 35,000/day. However, on August 9, the reception data of the scenic spot shows that the traffic volume of the scenic spot reached 50,000. On July 30th, Huangguoshu Scenic Spot has received 9,927,200 tourists this year, an increase of 36.1% over the same period of last year;

2018: From July 16th to August 14th, Huangguoshu Scenic Area received more than 30,000 visitors per day and 92,640 vehicles. In September, the total number of visitors received reached 970,107. Since the beginning of this year, there have been 3.4 million visitors to the Huangguoshu scenic spot.

2.1.4 Surrounding Environment

The area where Huangguoshu is located belongs to the mid-subtropical climate. It is mild throughout the year, and it is cool in winter and summer. The frost-free period can reach 290 days. During the period, the rainfall is abundant and the humidity is high. The relative humidity of the whole year is maintained at about 80%. The annual average temperature is 16.2 °C, the hottest month is 226 °C, and the annual time exceeds 30 °C for less than 10 days. It is mild in winter and cool in summer. The highest temperature is 35.3 °C and the lowest temperature is -8.5 °C. The

Table 1 Distance from Huangguoshu

Sight spot	Stone Village	Huangguoshu Hotel	Huashi whistle	Doupotang scenic spot	Huangguoshu Waterfall	Stalagmites hill	Daping land	Luoshitan Waterfall
Distance	4.5 km	1 km	7.8 km	1 km	2.5 km	312 km	175 km	4.8 km

Table 2 Distance from various scenic spots to the Tianxing Bridge

Sight spot	Hongyan monument	Tianxing hole	Longzhao tree	Red rock ancient Ficus	Yinlianzhuitan Waterfall
Distance	1.5 km	1.2 km	14.6 km	6.5 km	1.8 km

Table 3 Vegetation classification

Arbor	Bamboo woodland	Sparse woodland	Shrubbery	Unforested land	Planning for afforestation
2582.53	1	8.73	4566.47	734.2	660.87

Table 4 Classification by forest species

Protective forest	Special purpose forest	Timber forest	Economic forest	Firewood	Natural forest	Plantation
3228.73	1493.6	1129.8	326.87	946.53	4849.4	3043.47

annual temperature difference is small. 90% of the annual rainfall is concentrated in April–October.

The area of forest land in Huangguoshu scenic spot is 8553.8 hm² (the specific classification accounts for the following Tables 1 and 2). The proportion of middle-aged forest and near-mature forest is about 1666.67 hm², and the young forest is 690.67 hm²; the forest coverage rate of the whole district is 52.33%. Covering the surrounding tree area, the forest greening rate is 52.93%, and the forest stocking volume is 109,300 m³ [2].

The vegetation topography around the attraction is divided into two types according to different classification methods (Tables 3 and 4).

3 Technology Architecture

3.1 Pre-modeling Preparation

In the early stage of modeling, through the investigation of the overall situation of Huangguoshu scenic spot, and the related constraints, the constraints are selected according to relevant data, and the index system is constructed by considering the impact degree of attraction, self-attraction, surrounding environment and spatial factors; The standardization method of the standard deviation standardizes the indicators so that all the indicators can be transformed into data that can be uniformly compared and not restricted by the unit [3]. Finally, objectively and accurately reflect the actual influence degree of each indicator layer on the target layer, and the entropy

method is used to standardize each Indicators and definitions of the difference factor vector to achieve the purpose of determining the weight of each indicator [4].

3.2 Modeling Thinking

Firstly, the short-term forecasting method is used to predict the traffic volume of the updated scenic spots with corresponding observations; then the saturation of the scenic spot is used as an indicator to define whether the scenic spot is congested; then the model is combined with observing the changes of various scenic spots to predict the tourist traffic of the scenic spot, and realize fast [5]. The data is transmitted accurately, the tourists are positioned in real time and the tourists are induced to visit. Finally, considering the various costs between different scenic spots in the scenic spot, the A* algorithm [6] is used as the path search method to construct the road network model.

3.3 Solution Principle

The project model is based on science and operability. Therefore, the selected indicators of this project model are the most representative and most reasonable leading indicators, in order to achieve the purpose of combining the indicators to accurately describe the current situation in the scenic area. In addition, the methods and techniques used in the modeling process of this project are relatively advanced.

4 Technical Details

4.1 Design Thought

1. Collect and conduct quantitative comparative analysis of the current situation of scenic spots;
2. Extract the main limiting factors that affect the tourism route planning to conduct regional block division;
3. According to the weight ratio of index factors, use the improved algorithm to design the optimized route model.

4.2 *Research Technique*

Specific methods: the acquisition method was adopted to obtain the preliminary basic data, dazzled value method, short-term prediction method and A* path search algorithm for the later route planning and optimization. First investigation by the acquisition method related sites, get basic data for regional division, through the entropy method [7] to empowerment of index factors, on the basis of previous data and the observed value of scenic spots to short-term traffic forecasting, traffic update forecast to attractions, attractions passenger flow saturation as judgment index of scenic spot is crowded, finally USES the A* path search algorithm optimize the route takes time, provide more freedom to choose A reasonable route for tourists.

Study arrangement:

1. Two team members are responsible for data collection, inquiry and sorting of scenic spots, and early data collection;
2. Two team members predicted and analyzed the congestion degree and network route according to relevant data, and completed the literature review of scenic spot factors;
3. A team member made weight analysis of the predicted result route and each route factor, and assigned different weights to each scenic route;
4. The members of the whole group determine the feasibility of implementing the regional route based on the collection method and various practical data;
5. All members of the group sorted out the query data results and collected data, and completed the study.

4.3 *Feasibility Analysis*

1. The positioning of the research problem is feasible. The research on tourism development at home and abroad is still in a relatively lagging stage, often neglecting the influence of some related factors of seasonal hot scenic spots, resulting in scenic spot congestion, minimizing the utilization of resources, the level of demand and expectations of tourists can not meet the expected value and so on. This project takes this as the starting point to link the influencing factors of hot tourism season, the utilization rate of resources, the requirements of tourists level and route planning, and so on. The subjective and objective factors that affect the path planning and the research of path optimization are carried out in the network construction which is optimized in all aspects, and finally the route planning is optimized [8]. Therefore, the positioning of the research problem takes into account the imperfections of the current tourism development, and puts the theoretical research into practice to achieve a role in promoting.
2. The research method adopted in this project is feasible. The entropy method, short-term forecasting method and A* path search algorithm laid the foundation

for the project. The research and calculation method theory involved in this project have been successfully applied in the field of traffic path planning, and the existing research results have provided strong support for the project.

5 Prototype Testing

5.1 Constructing Index System

The hotspot scenic route planning depends on the influence of various influencing factors on its path selection, so it is necessary to calculate the weight of each influencing factor. The construction of the scenic spot index system mainly reflects the dominant thinking of multiple aspects of the consideration of the scenic spot as the carrier, and makes each measurement index reflect the most representative and most reasonable leading indicator among many indicators, so that the measurement index system can accurately describe The status quo in the scenic area, based on the scientific and operational behavior benchmarks, comprehensive consideration of the congestion level of the attractions, their own attractiveness, the surrounding environment and spatial factors as the impact indicators of the hotspot scenic route planning. Thus, getting the following Table 5.

5.1.1 Target Standardization

In the hot spot scenic spot impact index system, because the different index units have different properties, in order to remove the unit limitation of the data, it is transformed into dimensionless pure data, that is, the indicators are standardized so that all indicators can achieve unified comparison data. Here we use the standardization of the standardization to standardize the indicators.

Table 5 Hot spot scenic spot impact indicator system

Target layer	Criterion layer	Index layer
Comprehensive evaluation	Congestion	Passenger flow
	Spatial factors	Distance
	Attraction of Scenic Spots	Characteristic culture
		Number of attractors
	Surrounding Environment of Scenic Spots	Geographical conditions
		Public facilities

For the positive indexes, the upper limit result measurement formula is:

$$v = 1, 2, \dots, m; \quad i = 1, 2, \dots, n$$

For the negative indexes, the upper limit result measurement formula is:

$$v = 1, 2, \dots, m; \quad i = 1, 2, \dots, n$$

In the two inequalities, is the actual value of the v th evaluation unit on the i th indicator. is the maximum value of the i th indicator, is the minimum value of the i th indicator, $i = 1, 2, \dots, n$.

5.1.2 Confirm of Indexes Proportion

In order to objectively reflect the influence of each indicator layer on the target layer, the entropy method is used to determine the index weight. The procedure is as follows:

1. Standardize each indicator and calculating the proportion of Item i index of scenic spot v

$$p_{vi} = (X_{vi}) = \frac{X_{vi}}{\sum_{v=1}^m X_{vi}} = 1, \quad v = 1, 2, \dots, m, \quad i = 1, 2, \dots, n \quad (1)$$

Calculate the entropy of the i th indicator:

$$e_i = -k_v \sum_{v=1}^m (p_{vi} \ln(p_{vi})), \quad i = 1, 2, \dots, n \quad (2)$$

In the middle, $k > 0, e_i \geq 0$, The smaller the difference between a given i and X_{vi} , the larger e_i is. When X_{vi} is all equal, $e_i = 1$, then the index X_{vi} is meaningless compared with the scheme; When the difference between the index values of the schemes is larger, the smaller the e_i is, the greater the effect will be compared with the scheme. Due to $m = 5$, and so $= 0.6213$.

2. Calculate the difference coefficient of the index of the i th item, define the difference factor vector as $G = (g_1, g_2, g_3, \dots)$, the larger the g_i , the more important the indicator:

$$g_i = 1 - e_i, \quad i = 1, 2, \dots, n \quad (3)$$

3. Calculate the weight of each indicator by using the difference factor $g_i, W = (W_1, W_2, \dots, W_n)$:

$$W_i = g_v \sum_{i=1}^m g_v \tag{4}$$

According to the relevant data, the weight of each indicator is calculated as: = 0.418; = 0.325; = 0.0316; = 0.0409; = 0.0497; = 0.125.

5.2 Model

5.2.1 Short-Term Forecast of Passenger Traffic

We use short-term forecasting method to predict the passenger flow of the spot first. The so-called short-term forecast is based on the base forecast, and the visitor observation value is used to update the passenger flow in the corresponding time period. Since the opening time of the Huangguoshu scenic spot is 7:00–18:00, Therefore, we multiply the 9-h forecast by an update factor β , the attraction is v , and the short-term forecast for the time period $t + T$ is:

$$q_{vt+T}^{st} = q_{vt+T}^9 \cdot \beta \tag{5}$$

$$\beta = \left(\frac{\sum_{t=1}^t q_{vt}^{obs}}{\sum_{t=1}^t q_{vt}^9} \right)^{h(T)} \tag{6}$$

where q_{vt+T}^{st} is the predicted traffic for the v th attraction, β is the cost of the v attraction, $\sum_{t=1}^t q_{vt}^{obs}$ is the sum of the detected traffic of $t = 1$ to the current time period, $\sum_{t=1}^t q_{vt}^9$ is the sum of the predicted traffic of 9 h from $t = 1$ to the current time period, $h(T)$ indicates the degree of correlation of changes in passenger flow trends over several consecutive periods [9].

5.2.2 Establishment of Crowded Spots

The path optimization is related to the congestion degree of the attraction. To this end, the congestion degree of the scenic spot is defined first, and the saturation of the scenic spot is used as an index for defining whether the scenic spot is congested, that is,

$$\rho_v = \begin{cases} Q_v/C_v, & Q_v > C_v \\ 1, & Q_v \leq C_v \end{cases} \tag{7}$$

where is the number of visitors to the v th attraction, and v is the number of the attraction; is the maximum capacity of the attraction v . When the is larger, the greater the congestion of the attraction v .

5.2.3 Optimal Path Planning

Using the scenic spot to predict the passenger flow to realize the tourist induction, in the path search process, according to the scenic spot condition prediction model, the passenger flow status of each scenic spot is predicted within the entire scenic spot, and the predicted travel flow condition is used to calculate the travel cost of the tourist. In order to integrate the change of attractions from the starting time to the time of the attraction, the visitors can get the real-time update optimal path.

Using the A* algorithm as the path search method, since the model of the road network is constructed based on the idea of the scenic spot road, the scenic road nodes are no longer nodes in the traditional sense, but have a certain length, so the different scenic spots of the same scenic spot A* The valuation value of the algorithm heuristic information is not comparable, and the calculation method of the heuristic information valuation value needs to be improved, specifically: when implementing the update of the scenic road node, considering various costs between different scenic spots of the scenic spot. The improved A* algorithm examines the valuation function of the next node from the current node:

$$\begin{cases} f(i) = g(i) + h'(i) \\ h'(i) = h(i) + L \end{cases} \tag{8}$$

where $f(i)$ is the total estimate of node i ; $h'(i)$ is the improved heuristic information valuation value; $h(i)$ is the estimated cost of the best path from i to the target node; L represents the weighted sum of various costs caused by the distance between different attractions in the same scenic spot; $g(i)$ is the travel cost from the initial node to the i node after forecasting the traffic of the attraction, which can be expressed as:

$$g(i) = \sum_{l=1}^{l=i} (z(v) + s(v)) \tag{9}$$

$$z(v) = \rho v \cdot q_{dvt}^{st} \cdot \omega v \tag{10}$$

where is the congestion factor of the v th attraction, q_{vt+T}^{st} is the predicted traffic for the v th attraction, ω_v is the cost of the v attraction, $s(v)$ represents the sum of various weighted costs of the current attraction v to the next attraction $v + 1$, and the calculation method is:

$$s(v) = \sum_{k=1}^n \omega_k M_k \tag{11}$$

M_k represents the cost of the k th impedance, including the degree of congestion, time distance, number of attraction attractions, surrounding environment, etc., ω_k indicates its weight value. Finally, it is integrated by (5), (6), (7) and (10):

$$z(v) = \begin{cases} \frac{Q_v}{C_v} \cdot q_{vt+T}^9 \cdot \left(\frac{\sum_{t'=t-r}^t q_{vt'}^{obs}}{\sum_{t'=t-r}^t q_{vt'}^9} \right)^{h(T)} \cdot \omega_v, Q_v > C_v \\ q_{vt+T}^9 \cdot \left(\frac{\sum_{t'=t-r}^t q_{vt'}^{obs}}{\sum_{t'=t-r}^t q_{vt'}^9} \right)^{h(T)} \cdot \omega_v, Q_v \leq C_v \end{cases} \quad (12)$$

Flow chart of scenic passenger flow prediction path planning based on improved A* search algorithm is shown in Fig. 2.

Concrete implementation steps:

- Step 1: Select the initial attraction road node;
- Step 2: Query the road node of the scenic spot with which it can realize the switch;
- Step 3: Estimate the time from the starting point to each attraction T;
- Step 4: Forecasting tourist traffic after T time;
- Step 5: Calculate the degree of congestion and its corresponding cost;
- Step 6: Calculate the sum of the travel cost and the heuristic cost;
- Step 7: Take the least cost sight road as the next expansion node;
- Step 8: Determine if the attraction is the target attraction; yes, go to step 9; no, go to step 2;
- Step 9: Output optimal path.

6 Application Results

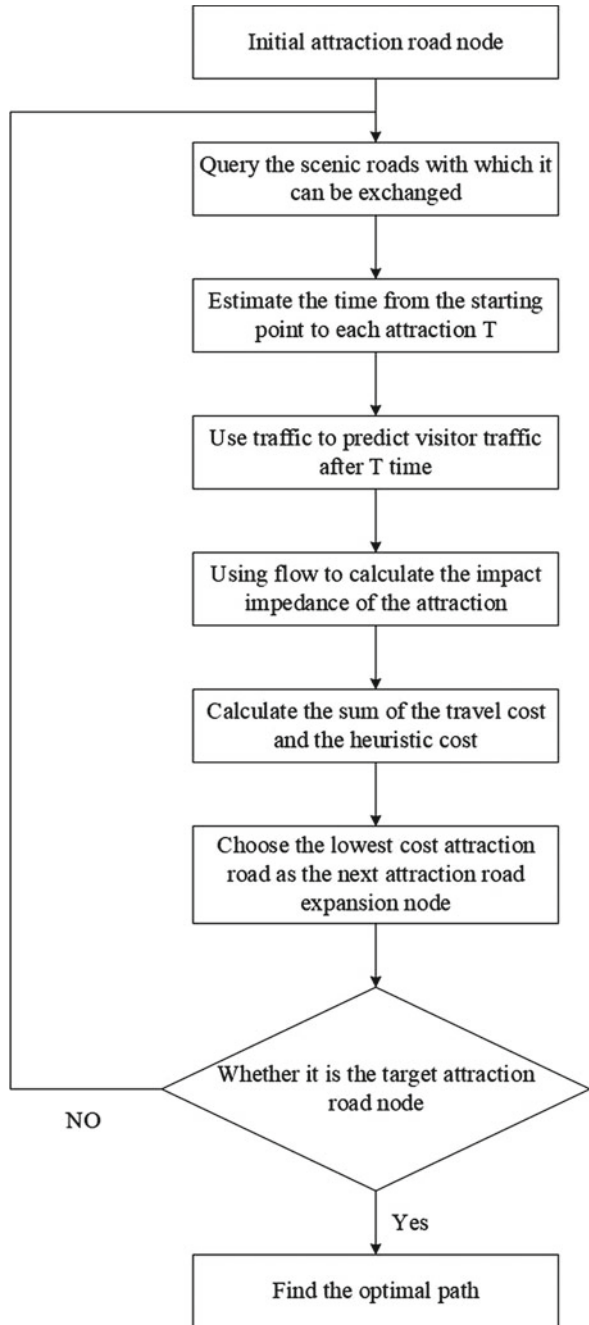
6.1 Model Solution

In order to improve the reliability of the experimental model, we selected Huangguoshu scenic spot as the research of route planning in this hot spot tourism scenic spot. The weight coefficients of each index and the forecast data of passenger flow of scenic spot are substituted into the improved A* search algorithm model. According to the process steps of optimal path planning, we can get a scenic spot in Huangguoshu. Real-time optimal path.

1. Setting up the initial tourist attractions:

Assuming that the Star Bridge scenic spot, that is, the Star Bridge as the starting point of the scenic spot.

Fig. 2 Scenic passenger flow forecast path planning flow chart



2. Through scenic spots:

According to the guide map of Huangguoshu Scenic Area, starting from the Tianxing Bridge (0) and considering the time factors, we can choose Hongyan Stele (1), Tianxing Start Cave Cave (2), Longchaoshu (3), Hongyan Gurong (4) and Silver Chain Falling Pool Waterfall (5).

3. Calculating the time T from the starting point to each scenic spot:

According to the distance data of each scenic spot inquired, supposing that tourists choose walking mode to visit, calculating the walking speed of tourists according to 1 m/s, the following results are obtained:

$$T_1 = 25 \text{ min}; T_2 = 20 \text{ min}; T_3 = 243.3 \text{ min}; T_4 = 108.3 \text{ min}; T_5 = 30 \text{ min}.$$

4. Prediction of tourist flow in scenic spots after T-time:

Because in the example, we assume that tourists choose the way of walking, for relatively distant scenic spots, tourists have a longer walking speed. Therefore, when predicting passenger flow, two scenic spots, Longchaoshu (3) and Hongyan Gurong (4), are not considered. In reality, when the scenic spots are relatively dispersed, various ways should be considered and used for sightseeing, such as setting up a tour inside the scenic area. Car viewing, parking and sharing bicycles, etc. Assuming that in the period from 8:00 to 9:00, then $t = 2$ and $T = 1$, according to the predicted passenger flow data of the total passenger flow detected and the opening time of the scenic spot for 9 h on the same day, the formula (5) and (6) are substituted to calculate:

$$\begin{aligned} &0.384 \\ &= 917.376; = 636.288; = 705.024 \end{aligned}$$

5. Optimal path calculation:

The model (7) is used to judge the congestion degree of each scenic spot after T-time. According to the forecast result of passenger flow, Judging whether the passenger flow of Hongyan Stele (1), Tianxing Cave (2) and Silver Chain Falling Pond Falls (5) after T-time exceeds the maximum capacity of the three scenic spots, some of the data could not be found clearly. We assume that the passenger flow of Hongyan Stele (1) scenic spot is larger than the capacity of scenic spot after T-time, and neither Star Cave (2) nor Silver Chain Fall Pool (5) exceeds the maximum capacity. The results are as follows: $= 1.31; = = 1$. Based on the investigation of the travel evaluation and heuristic evaluation of the next scenic spot, a judgment is made from the Tianxing Bridge (0) to the next scenic spot, that is, the Tainting Bridge—Tainting Cave—Silver Chain Falling Pool Waterfall; When a new scenic spot is reached, the scenic spot can be regarded as the starting point of the path, and the updated optimal path can be obtained by recycling and recalculating.

6.2 Result Analysis

Aiming at the most path planning of hot tourist attractions, this paper chooses Tianxing Bridge scenic spot to study and verify the reliability of the model. According to the past data and taking the Tianxing Bridge as the starting point, this paper predicts the passenger flow of the Tianxing Bridge scenic spot from 8:00 to 9:00. In order to simplify the calculation, this paper chooses a single mode of transportation, assuming that the tourists choose to walk for the tour, so for some scenic spots that are a little farther away, the walking time is a little longer. In reality, for the more scattered scenic spots, sightseeing cars and shared bicycles are often arranged between them for sightseeing. The crowding degree of scenic spots in T-time largely affects the choice of paths. By judging the crowding degree and cost of each scenic spot after T-time, the Star Bridge is finally the starting point. Optimal path: Tianxing Bridge to Tianxing Cave to Silver Chain Falls to Red Rock Stele. The purpose of this model method is to update it in real time according to different conditions. When you arrive at the next scenic spot, you can recalculate it and get the latest route. This method not only takes into account many factors, but also keeps updating the route, so that tourists can save time and better arrange the tour time in the process of touring.

7 Value Orientation

1. Social value:

The reasonable arrangement of the resources of each scenic spot. Real-time detection of crowd congestion in scenic spots. The aim of the sustainable development of the scenic spot is to disperse the crowd in time, maximize the efficiency of comprehensive utilization of the scenic spot resources and the efficiency of natural ecology protection, and finally evacuate the crowd. The collection method and the search method are used to obtain the preliminary basic data, and the improved A* Path search algorithm is used to carry out the later route planning and optimization. Optimize the time spent on the route, provide more reasonable routes for tourists to choose freely. Reduce the waiting time of tourists in sightseeing, and get the maximum value in the effective time.

2. Academic value:

At present, more and more people will choose to go out in the hot tourist season to see the magnificent mountains and rivers of the motherland. However, the contradiction between the limited resources of scenic spots and the increasing number of tourists becomes more and more prominent, and when unexpected incidents such as the unreasonable arrangement of emergency sightseeing buses in scenic spots and the failure of tourists' route scheduling occur, it will make the utilization ratio of scenic spot resources decrease greatly, and then lead to the uneven distribution of tourists and the phenomenon of structural congestion. People need to solve this contradiction.

However, domestic scholars and related tourism industry have little research on the structural congestion and other related issues, so there is no effective program, so this project has great prospects for development.

3. Commercial value:

If the project is successful, it will fundamentally change the traditional way of thinking of tourists and make it take on the task of evacuating tourists from hot spots to relieve structural congestion. At the same time, the scenic spot planning agency, namely the Management Committee, can optimize the layout of the scenic spot structure in a timely manner based on the latest tourist data provided by the huge database behind the technology. To promote the growth of tourism consumption in scenic spots and to create greater economic benefits. And we have not found similar experience before it, the original technology to fill the gap in this field, in the same enterprise lead, mission and high status.

8 Innovative Features

1. The index is standardized based on the method of range standardization, and the index is transformed into the data which is not influenced by the unit.
2. Entropy method is used to determine the weight of each index and define the Coefficient of difference, which objectively reflects the influence of each index layer on the target layer.
3. The short-term forecasting method is used to forecast the tourist flow of the scenic spot in real time, so that the tourist can update the route in time.
4. Considering the length of the scenic spot road nodes, the improved A* Algorithm is used as the path search method to update the route and construct the road network model.

9 Conclusion

By consulting relevant literature, multi-channel investigation and data collection, this paper has gained a certain insight into the planning of tourist routes in hot spots. As the route planning in hot spots is a multi-factor problem, which is affected by many indicators, this paper uses the method of entropy to assign the corresponding weight to each index factor. The improved A* path search algorithm is used to plan the optimal path, considering the influence of various factors and updating the route in time, which saves tourists' time and improves tourists' satisfaction in the process of touring. With the aid of this investigation and research, the overcrowding in the hot scenic spots is expected to be relieved to some extent. In the future, we will continue to pay attention to this problem, giving full play to the spirit of college students

who are good at finding and solving problems, and making continuous progress and innovation.

References

1. Wu Y, Yang X (2014) A review of the impact of tourism congestion
2. Luo R, Feng K, Li S (2018) Ecological environment quality assessment of Huangguoshu scenic spot (6):131
3. Sara D, Scott M (2001) *Ann Tourism Res*
4. Cai W (2009) The role of operational research in tourism route planning
5. Terry VL, *Tourism Management*
6. Zou S, Ruan J, Liu B (2008) Application of shortest path algorithm in tourism route planning
7. Zhou Y, Yang S, Xian C (2014) Comprehensive benefit evaluation of land use based on entropy method. *Guangzhou Agric Sci* (24):181
8. Cao Z, Han F, Wang Y (2012) A new dynamic path optimization method for urban traffic network. *Acta Electron Sin* (10):2064–2065
9. Zeng J, Wu Q, Qiu G, Guo C (2011) Research and implementation of real-time information transmission system for intelligent bus operation and internal congestion

Research on Energy-Saving Driving for Transport Vehicles Considering Actual Load



Chunming Li, Tao Zhang, Xiaoxia Sun, Chunming Shao, and Guozhu Wang

Abstract To improve the economy performance of plug-in hybrid electric transport vehicle (PHETV) on urban roads, a pseudo-spectral based speed planning method is proposed for connected vehicles to assist driver to pass continuous intersections with fewer stop-and-go behaviors. Firstly, the effective green light time is estimated based on the traffic flow obtained by V2I. Secondly, the actual vehicle load is identified through neural network which is trained by the historical data. Finally, the time-domain optimal control problem with piecewise head and tail constraints is constructed and solved by the pseudo-spectral method to minimize the travelling energy consumption. Simulation results show that the proposed control strategy can make fuel consumption of PHETV reduce by 6.71% compared with traditional method.

Keywords Hybrid electric vehicle · Energy saving · Pseudo-spectral method

1 Introduction

The application of car networking technology is revolutionizing the development of the transportation industry. By using vehicle-to-infrastructure (V2I) wireless communication [1, 2], traffic jams or unnecessary stop-and-go behavior in the main road are reduced to reduce fuel consumption is the important research for current urban road energy-saving driving. Compared with the eco-driving plan of ordinary cars, urban transport vehicle are also suitable carriers for the application of ecological driving strategies. On the one hand, the transport vehicle has a heavy load and is mainly driven under urban roads with traffic signal restrictions. Passive stop-and-go behavior leads to excessive fuel consumption [3]. On the other hand, due to the all-weather delivery of goods, it is necessary to ensure a reliable source of vehicle power. Compared to pure electric vehicle, the Plug-in hybrid electric drive configuration has potential to improve the fuel economy and the maneuverability, and bring more reliable onboard electricity supply [4]. Therefore, the plug-in hybrid electric

C. Li · T. Zhang (✉) · X. Sun · C. Shao · G. Wang
China North Vehicle Research Institute, Beijing 100072, India
e-mail: ztao1208@126.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_38

509

transport vehicle (PHETV) with connected and automated technology will play an important role in urban transport distribution in the future.

The energy management strategies (EMSs) of most hybrid electric vehicles (HEVs) can be divided into rule-based and optimization-based management. Rule-based energy management strategies are widely used in the automotive industry. In comparison, the optimized energy management strategy has better theoretical fuel economy, such as dynamic programming (DP) [5], Pontiac's minimum principle (PMP) [6], Equivalent Consumption Minimization Strategy (ECMS) [7], Model Predictive Control (MPC) [8] and Reinforced Learning (RL) [9]. These methods are often based on certain specific driving conditions for off-line optimization control strategies, which is too ideal for actual driving situations. Typically, for applications in HEVs on urban roads, traffic information (e.g. traffic signal phase and timing), as well as road conditions (e.g., speed limits), is considered to optimize EMSs.

The nonlinearity and uncertainty of vehicle dynamics, the limited computing resources of the control unit, the time-varying road/traffic conditions, etc., make the problem of hybrid energy management optimization under dynamic traffic more challenging. Due to the cyclical changes of traffic lights, the speed of the vehicle is strictly limited by the signal time. Traffic queues causes the green light time for vehicles to pass less than the maximum green signal period provided by signal phase and timing (SPaT), and this makes optimization more challenging. However, most of the research on the energy management of PHEV in urban roads does not take into account the constraints of actual traffic queuing. At the same time, due to the uncertain driving load of the transport vehicle and the large range of weight variation, it has a great impact on the results of system energy saving optimization. How to accurately estimate the actual load of the transport vehicle based on the vehicle driving data also needs to be studied and introduced into the EMS.

As a typical direct method classified as an nonlinear programming (NLP) method, pseudo-spectral methods have been increasingly used for numerical solving of optimal control problems for various dynamic systems [10, 11]. The pseudo-spectral method is easy to deal with the segmentation problem, which can transform the continuous intersection traffic problem into a continuous single-signal intersection speed planning problem. By splicing the constraints of multi-segment planning, the energy management problem of continuous intersections is solved with the goal of the lowest global energy consumption.

Inspired by the above problems, this paper uses the deep learning network to identify the actual load of transport vehicles, and proposes the pseudo-spectral method to solve the energy-saving traffic problem of multi-section by constructing the effective signal lamp model of intersection.

2 Control-Oriented Model of the PHETV

The vehicle used in this study is a modified PHETV. The topology diagram of the real vehicle and its structure is shown in Fig. 1. The powertrain consists of two power

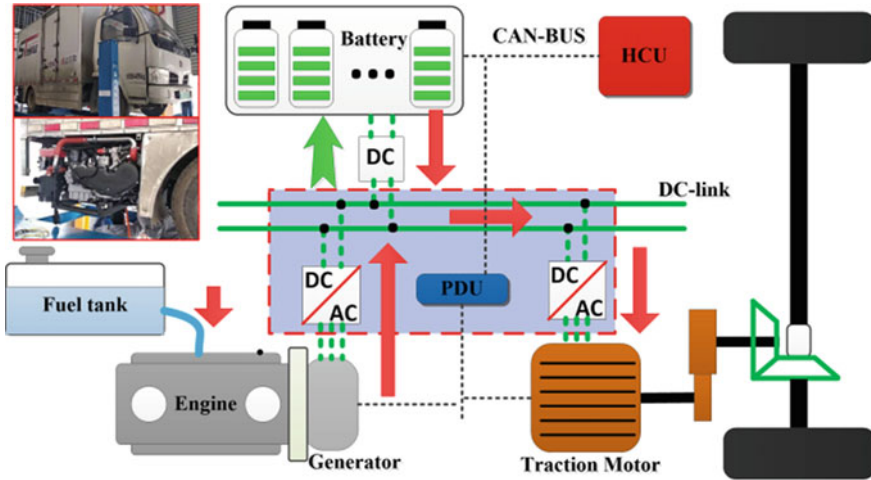


Fig. 1 PHETV and power topology diagram

sources, namely an EGS set (a diesel engine coupled to a generator) and a lithium-battery package. The power demand of the vehicle may be supplied by any of the two power sources, or the battery can be charged via the power distribution unit (PDU).

2.1 Power Balance of Powertrain

When the vehicle enters the main road, the state vector $\mathbf{x}(t)$ is defined as:

$$\mathbf{x}(t) = [x_1(t), x_2(t)]^T = [s(t), v(t)]^T \tag{1}$$

where $s(t)$ the distance from the arterial entrance s_0 , and $v(t)$ is the speed of the vehicle. The relationship between vehicle position, velocity and acceleration is as follows:

$$\begin{aligned} \dot{s}(t) &= v(t), \\ \dot{v}(t) &= \frac{T_{mot} i_{gr} \eta_{trac}}{M r} - g f_r \cos \alpha - \frac{1}{2M} \rho A_f C_D v^2 - g \sin \alpha \end{aligned} \tag{2}$$

where T_{mot} is the traction force of the driving motor, η_{trac} is the transfer efficiency of traction, i_{gr} represents the transmission ratio between the driveline and the traction motor, g is the acceleration of gravity, f_r is the rolling resistance coefficient, M is the vehicle equivalent mass, C_D is the aerodynamic drag coefficient, A_f is the frontal area, ρ is the air density, α is the road gradient, r is the radius of the tire. Vehicle

dynamics is defined as $\dot{\mathbf{x}}(t) = f[\mathbf{x}(t), \mathbf{u}(t)]$, and the control variable vector $\mathbf{u}(t)$ is the acceleration rate at each time.

Based on the tire rotational speed w_o and the driveline power demand P_m , the torque T_{mot} , speed w_{mot} , and the power balance of powertrain regarding the traction motor can be expressed as:

$$\begin{cases} P_{gen} + P_b = [P_m \eta_m(T_{mot}, w_{mot})^{-\text{sign}(P_m)} \eta_b^{-\text{sign}(P_m)}] + P_a \\ T_{mot} = \frac{P_m}{w_o i_{gr}}, \quad w_{mot} = w_o i_{gr} \end{cases} \quad (3)$$

where i_{gr} represents the transmission ratio between the driveline and the traction motor, P_{gen} is the generator power, P_m is the required power of the transmission system, P_b is the output power of the battery, P_a is the sum of auxiliary power, which can be regarded as the fixed value. η_m is the motor efficiency, η_b is the cell efficiency, and $\text{sign}()$ is the signum function.

2.2 EGS Model and Battery Model

In the PHEV system, the output voltage, output power, torque and speed of the engine-generator set meet specific relationships. In this study, the EGS Model and the first-order internal resistance Battery model were directly adopted by Wei et al. [10], Guo et al. [12]. In addition, battery life models can be further introduced to enhance coordination between fuel economy and battery life.

3 Vehicle Load Identification Model and Traffic Model

3.1 Identification of Vehicle Load

During driving, the vehicle system dynamics involve lots of physical variables and greater random disturbances, which makes it not easy to directly calculate the vehicle mass under dynamic conditions. Therefore, a novel approach based on deep learning algorithm is proposed and adopted to predictive the vehicle mass by collected driving data. The model predictive methods can be divided into sequential algorithms and non-sequential algorithms. The relationship among sequences of drive data is time-dependent to some extent, thus the DNN (deep neural network) algorithm is suitable for this case and adopted to accurately predict the vehicle mass load here [13]. The essence of DNN is to learn the features among the data by building a machine learning model with hidden layers and a large amount of training data, and through the feature transformation layer by layer, make use of rich feature nonlinear combination, so as to achieve the improvement of prediction accuracy.

Table 1 The relevant network test results

Parameter	Value
Single training time	0.5 s
Training number	1080
Training MSE	0.021
Convergence speed	9020
Batch inference speed	1.1 (100 times /s)

In the experiment, the driving data of acceleration and deceleration under different load conditions were collected, including: single-cycle rate of change of brake pedal, depth of brake pedal, single-cycle rate of change of accelerator pedal, depth of accelerator pedal, speed, timing sequence, load and other information. The data sampling frequency is 50 Hz. In general, the vehicle can complete acceleration or deceleration operations within 4 s, so the effective data is segmented into each 80 pieces (i.e., 4 s interval) data into a data slice, and the effective cleaning, normalization, de-timing and other processing are also carried out.

The data can be sorted as $\begin{bmatrix} X_1 & X_2 & \dots & X_n \\ y_1 & y_2 & \dots & y_n \end{bmatrix}$, where $[X_1 \ X_2 \ \dots \ X_n]$ is the inputs, i.e. pedal cycle change rate, time, speed, etc. $[y_1 \ y_2 \ \dots \ y_n]$ is the outputs corresponding to the vehicle load data, and n is the index number. The constructed DNN includes three layers: input layer, hidden layer and output layer. The parameters are set as total parameters of 692,565, trainable parameters of 689,837 and non-trainable parameters of 2728. The relevant network test results are shown in Table 1.

From test results, since the DNN network has the advantages of relatively simple structure and low consumption of calculation resources, it is easy to be applied into embedded devices and good for practical engineering applications. To observe the effect of network estimation, 200 groups of training data fragments were randomly selected, and the predicted results were shown in Fig. 2. The horizontal axis was the data number, and the vertical axis was the normalized value of load.

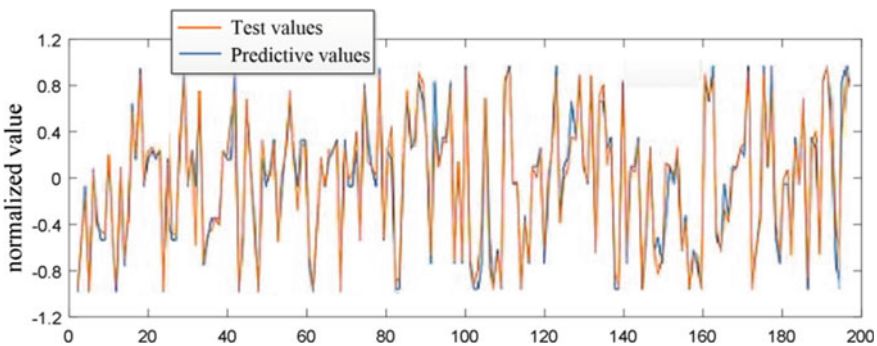


Fig. 2 Prediction results of training data fragments

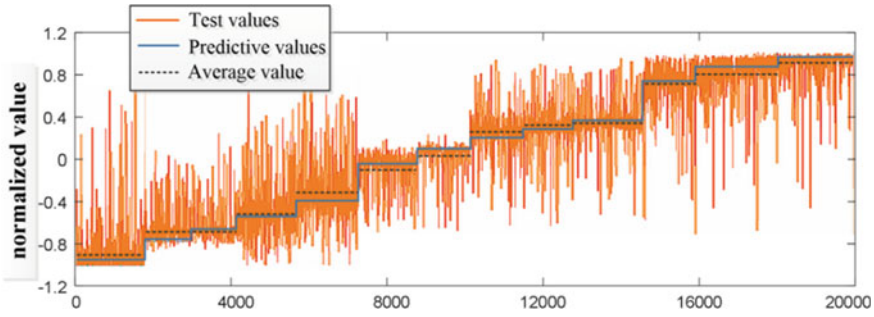


Fig. 3 Prediction results of continuous data

The results show that there is the errors of vehicle load between the real and predictive values by DNN network, but they are small, and the change trend of the value is consistent. To further observe the generalization ability of DNN, we extracted a total of 100,000 sets of multi-stage continuous driving data within the full load range of 2.9t–4.5t, and the predicted results are shown in Fig. 3.

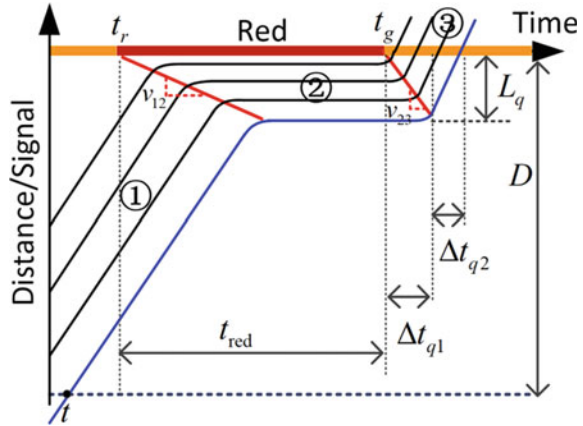
The results show that the predicted results fluctuate, and the statistical MSE error is 0.0337. This is because the actual driving conditions are complex and changeable, and the generalization ability of the network in the face of massive data is general, so the credibility of the single prediction result is low. But, overall, the DNN network has already possessed certain estimation ability. On the other hand, the results show that within the same load range, the average value of the predicted value (red dotted line) is close to the actual value. Therefore, in practical application, the vehicle load can be obtained by using the method of continuous prediction and averaging within the range.

3.2 Effective Signal Light Model

In the networked city traffic scenario, the traffic flow q , density ρ , and average vehicle speed v of different road sections and different time periods can be obtained by V2I. Since the traffic flow changes little in a short period of time, the short-term future traffic flow can be approximately replaced by the nearby historical data. Figure 4 shows the movements of a series of vehicles approaching and passing an intersection with traffic waves. According to the literature [14], the length of parking queue L_q at the intersection can be approximately estimated as:

$$L_q = \begin{cases} \frac{v_{12}}{v_0 + v_{12}} [D - v_0(t_r - t)], & \forall t \in \left[t_r - \frac{D}{v_0}, t_g + \frac{v_{12}(t_g - t_r)}{v_{12} + v_{23}} \right] \\ 0, & otherwise \end{cases} \quad (4)$$

Fig. 4 Traffic dynamics at an intersection



Here, t_r and t_g are the times that the traffic signal turns to a red and green indication, respectively. D is the distance to the intersection, v_0 is the initial speed of host vehicle. The deceleration wave v_{12} and the acceleration wave v_{23} can be obtained by the following formula [15]:

$$v_{12} = \frac{q_1}{\rho_1 - \rho_2}, \quad v_{23} = \frac{q_3}{\rho_3 - \rho_2} \tag{5}$$

where q_1 and q_3 are the traffic on roads stage 1 and 3 respectively. ρ_1 , ρ_2 and ρ_3 are the traffic density on roads stage 1, 2, and 3 respectively.

According to parking queue length and accelerating wave speed, the impassable time for host vehicle during the green signal time is calculated, and the effective red light signal time t_f is gained, which is the end time of the traffic queue.

$$t_f = t_{red} + \Delta t_{q1} + \Delta t_{q2} = t_{red} + \frac{L_q}{v_{23}} + \frac{L_q}{v_0} \tag{6}$$

4 Construction of Optimization Problem

Suppose that the planned road has n signalized intersections, the current time and position information of vehicles are expressed as $\mathbf{A}_I = (\tilde{T}_i^A, \tilde{D}_i^A)$. The time \tilde{T}_i^B and position \tilde{D}_i^B of vehicles arriving at the nearest intersection can be estimated by traffic flow and SPaT information. For an intersection i , an optimal control model for minimizing fuel consumption is established:

$$\min_{a(t)} \Gamma(i) = \int_{t_0^i}^{t_f^i} F(v(t), a(t)) dt \quad (7)$$

subject to the state constraints:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= f(\mathbf{x}(t), \mathbf{u}(t)), \quad t_0^i \leq t \leq t_f^i \\ \mathbf{x}(t_0^i) &= \mathbf{x}_0^i, \quad \forall i, \\ \mathbf{x}(t_f^i) &= \mathbf{x}_f^i, \quad \forall i, \end{aligned} \quad (8)$$

where \mathbf{x}_0^i is the vehicle state when it enters intersection i ; \mathbf{x}_f^i is the vehicle state when it leaves intersection i . queue limitation constraint:

$$\begin{aligned} s(\tilde{T}^A) &\leq \tilde{D}_i^A, \quad \forall i, \\ s(t_f^i) &= \tilde{D}_i^B, \quad \forall i, \\ t_f^i &\leq \tilde{T}_i^B, \quad \forall i, \end{aligned} \quad (9)$$

and define the lower and upper bounds for state and control variables:

$$0 \leq v(t) \leq v_m, \quad a_m \leq a(t) \leq a_M, \quad \forall t_0^i \leq t \leq t_f^i, \quad (10)$$

where v_M and a_M denotes maximal speed and maximal acceleration rate respectively. a_M is the maximal deceleration rate.

The traffic problem of continuous road is regarded as the optimal control problem of multi-stage single. And a multi-stage optimal control problem can be constructed by introducing the linkage constraints between boundary states of adjacent control stages, as follows:

$$\min_{a(t)} \Gamma = \sum_i \Gamma(t) = \sum_i \int_{t_0^i}^{t_f^i} F(v(t), a(t)) dt \quad (11)$$

Except satisfying the formulas 8, 9 and 10, it is also subject to:

$$\begin{aligned} \mathbf{x}(t_0^i) &= \mathbf{x}_f^{i-1}, \quad \forall i, \\ \mathbf{x}(0) &= \mathbf{x}_0, \\ t_0^1 &= 0, \\ s(0) &= 0, \\ s(t_f^n) &= L, \end{aligned} \quad (12)$$

where $\mathbf{x}(t_0^i) = \mathbf{x}_f^{i-1}$ represents the linkages between adjacent stages. Note that, the terminal times t_f^i are decision variables, including the total travel time $T = t_f^n$. Therefore, this is an optimal control problem of free termination time, which can be solved numerically by the gradient-based method offered by Pseudo-spectral Optimal Control Software (POPS) [16].

5 Numerical Examples

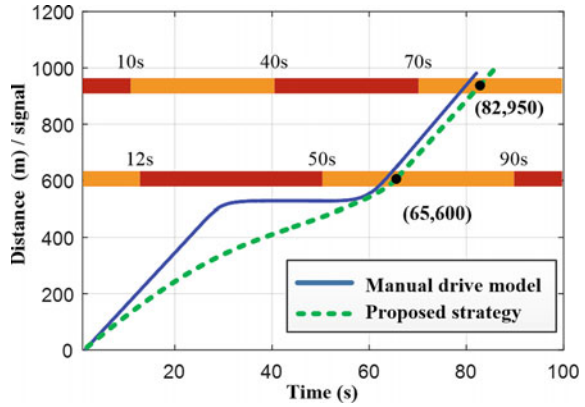
A simulation scene of four intersections is built, and the road information is shown in Table 2. In the simulation, assuming the initial vehicle speed $v_0 = 65$ km/h, the minimum road traffic speed is $V_{\min} = 25$ km/h, the maximum speed limit is $V_{\max} = 65$ km/h. The turning scenario is ignored for the moment, the vehicle is assumed to be empty with $M = 2.9t$.

Considering only two signal intersections, the optimization results are shown in Fig. 5. In order to compare the optimization effect, the manual driving model with fixed speed is introduced in the simulation [17]. According to the calculation result, the first intersection queue dissipation point coordinates (65, 600) is obtained, which means that the formation of the queue at the first intersection will dissipate in 65 s, 600 m. From the manual driving curve (blue solid line), it can be seen that if the energy-saving driving speed is not controlled, the vehicles will enter the queue at the first intersection, and then stop and wait until they follow the queue through the intersection. On the contrary, the energy-saving driving track (green dotted line) obtained by the Pseudo-spectral method obviously adjusts the speed actively to ensure that the vehicle following at the end of the queue can pass the first one intersection without stopping. Similarly, the coordinate of dissipation point of the second section of the road queue (82, 950) is estimated. Due to the short remaining distance and sufficient green signal time, the optimized speed curve of energy-saving driving is

Table 2 Multi-intersection traffic information

Parameter	Values			
Traffic light number	1	2	3	4
Distance (m)	600	950	1450	1800
Initial red phase (s)	50	10	44	16
Red/green light time (s)	38/40	30/30	36/36	30/30
Traffic q_1 (veh/h)	1000	850	900	700
Traffic ρ_1 (veh/km)	16	16	15	14
Traffic q_3 (veh/h)	1400	1400	1400	1300
Traffic ρ_3 (veh/km)	20	22	24	20
Traffic ρ_2 (veh/km)	120	120	110	90

Fig. 5 Vehicle location trajectory at two signalized intersection



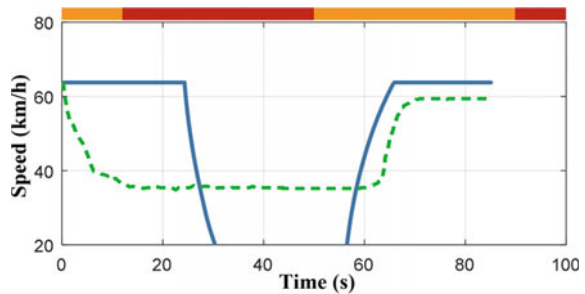
close to the manual driving track, and all of the vehicles pass the second intersection smoothly.

The speed curve is shown in Fig. 6. The recommended speed for optimization is less than the initial speed (65 km/h), and the vehicle gradually reduces to the appropriate speed. After passing the first intersection, the second section is driven at high speed. If the queue is not considered, vehicles will stop at the first intersection, which hinders the effectiveness of energy-saving system in reducing fuel consumption.

The simulation results of four complete intersections are shown in Fig. 7. From the simulation results, we can see that in the process of manual driving (blue solid line) through the four intersections, there are two parking processes due to the traffic flow queuing, and the speed trajectory of energy-saving driving (green dotted line) passes through the intersection smoothly. In addition, for the energy-saving driving trajectory, when passing through the fourth intersection, the estimated queue dissipation point coordinates are (142, 1800), but the optimized speed trajectory lags behind the dissipation point. This is because the vehicle arrives at the third intersection late and the planned speed exceeds the upper limit of safe speed, so it can only drive at the highest speed to approach the fourth intersection.

For the sake of comparison, the charge-depleting/charge-sustaining (CD/CS) method [18] and DP algorithm [5] are also taken into account in the simulation,

Fig. 6 The speed trajectory



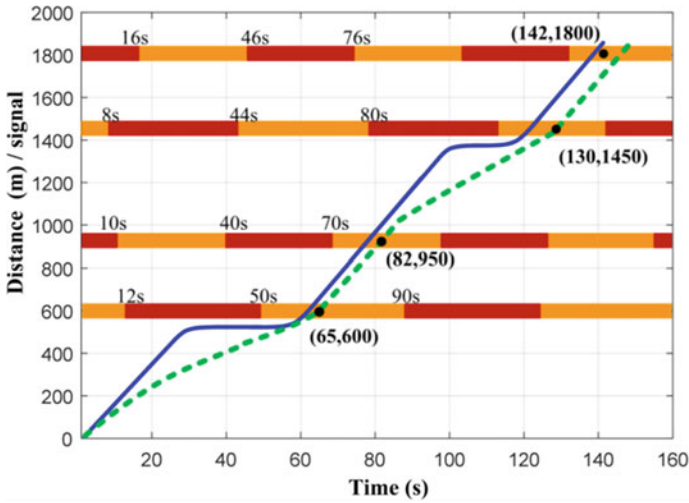
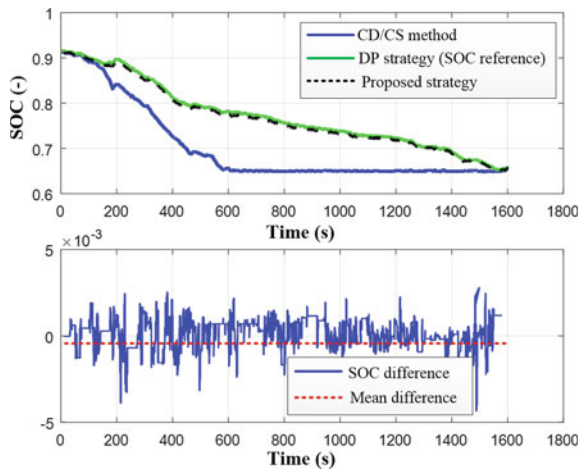


Fig. 7 Vehicle location trajectory under four intersections

Fig. 8 SOC tracking performance



and the initial SOC is set as 0.91. In order to observe the numerical change of energy consumption, the simulation cycle was executed 10 times. Figure 8 depicts the results of SOC profiles under different strategy. Result shows that the proposed strategy can effectively track the reference curve with a small error, and the tracking difference is restricted in the range of -0.005 to 0.005 , whose mean value is only -8.658×10^{-4} yielding the effectiveness of the terminal cost in pseudo-spectral method controller. Table 3 lists the equivalent cost results of RMB under different control algorithms. Compared with the CD/CS method, the DP algorithm and the pseudo-spectral method can reduce the total cost by 7.97% and 6.71% respectively. The difference may be

Table 3 Cost results by CD/CS, DP, and proposed strategy

Method	Fuel cost (CNY)	Electricity (CNY)	Total cost (CNY)	Cost saving (%)
CD/CS	3.47	1.37	4.77	–
DP	3.04	1.37	4.39	7.97
PSM	3.25	1.38	4.45	6.71

due to the different “grid number” of state variables used by the two algorithms, In spite of the differences, it can still be considered that the optimization strategy based on pseudo-spectral method can achieve the optimization effect of DP.

6 Conclusion

In this paper, a connected control algorithm for energy saving driving on urban roads is proposed for plug-in hybrid electric transport vehicle. The actual vehicle weight is identified through DNN network, and the traffic information is used to estimate the effective green light time at intersections. Based on the above information, the vehicle energy saving trajectory at continuous intersections is solved by using the pseudo-spectral method. Simulation results show that the proposed strategy can achieve the energy saving effect of dynamic programming algorithm. Compared with the CD/CS method, the pseudo-spectral method can reduce the total cost up to 6.71%.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 51705480).

References

1. Guo N, Zhang X, Zou Y et al (2020) A computationally efficient path following control strategy of autonomous electric vehicles with yaw motion stabilization. *IEEE Trans Transp Electrif* 6(2):728–739
2. Katsaros K, Kernchen R, Dianati M et al (2011) Application of vehicular communications for improving the efficiency of traffic in urban areas. *Wirel Commun Mob Comput* 11(12):1657–1667
3. Butakov VA, Ioannou P (2016) Personalized driver assistance for signalized intersections using V2I communication. *IEEE Trans Intell Transp Syst* 17(7):1910–1919
4. Zhang F, Hu X, Langari R et al (2019) Energy management strategies of connected HEVs and PHEVs: recent progress and outlook. *Progress Energy Combust Sci* 73(July):235–256
5. Chen Z, Mi CC, Xu J et al (2014) Energy management for a power-split plug-in hybrid electric vehicle based on dynamic programming and neural networks. *IEEE Trans Veh Technol* 63(4):1567–1580
6. Nguyen BH, German R, Trovao JPF et al (2019) Real-time energy management of battery/supercapacitor electric vehicles based on an adaptation of Pontryagin’s minimum principle. *IEEE Trans Veh Technol* 1:1–1

7. Zeng Y, Sheng J, Li M (2018) Adaptive real-time energy management strategy for plug-in hybrid electric vehicle based on simplified-ECMS and a novel driving pattern recognition method. *Math Probl Eng* 2018:1–12
8. Li XM, Han LJ, Liu H et al (2019) Real-time optimal energy management strategy for a dualmode power-split hybrid electric vehicle based on an explicit model predictive control algorithm. *Energy* 172:1161–1178
9. Wu Y, Tan H, Peng J et al (2019) Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus. *Energy* 247:454–466
10. Wei S, Zou Y, Sun F et al (2016) A pseudo-spectral method for solving optimal control problem of a hybrid tracked vehicle. *Appl Energy* 194:588–595
11. Xu S, Deng K, Li SE et al (2014) Legendre pseudo-spectral computation of optimal speed profiles for vehicle eco-driving system. In: 2014 IEEE intelligent vehicles symposium proceedings, pp 1103–1108
12. Guo N, Zhang X, Zou Y et al (2019) Computationally efficient nonlinear model predictive controller for energy management of tracked hybrid electric vehicles. 2019 IEEE VPPC, Hanoi, Vietnam, pp 1–6. <https://doi.org/10.1109/VPPC46532.2019.8952217>
13. Zhang J, Zheng Y, Qi D et al (2016) DNN-based prediction model for spatio-temporal data. In: The 24th ACM SIGSPATIAL international conference, (92), pp 1–4. <https://doi.org/10.1145/2996913.2997016>
14. Zhang T, Zou Y, Zhang X et al (2020) Predictive eco-driving application considering real-world traffic flow. *IEEE Access* 99:1–1. <https://doi.org/10.1109/ACCESS.2020.2991538>
15. Lighthill MJ, Whitham GB (1955) On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proc R Soc London, Ser A, Math Phys Sci* 1178(229):317–345
16. Rao AV, Benson DA, Darby CL et al (2010) Algorithm 902: GPOPS, A MATLAB software for solving multiple-phase optimal control problems using the gauss pseudospectral method. *ACM Trans Math Software* 37(2):22
17. Kesting A, Treiber M, Helbing D (2010) Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. *Phil Trans R Soc A Math Phys Eng Sci* 1928(368):4585–4605
18. Chen Z, Guo N, Shen J, Xiao R, Dong P (2018) A hierarchical energy management strategy for power-split plug-in hybrid electric vehicles considering velocity prediction. *IEEE Access* 6:33261–33274

Research on the Coupling Relationship Between Passengers and Vehicles Based on Anylogic Simulation



Jiahui Li, Siyuan Hao, Chengwu Jiao, Nale Zhao, and Keman Wu

Abstract With development of the urbanization process, more and more integrated passenger transport hubs were constructed. Some softwares are used to simulate traffic flow around the hub to plan traffic form layout and simulate the passengers in the hub to optimize facilities layout and walk access. In fact, passenger transport hub is the connection between passengers and vehicles, vehicles pick-up passengers in hubs, and passengers can transfer different means of transportation in hubs. But there are few simulation researches on the coupling relationship between passengers and vehicles. In fact, the poor connection between pedestrians and vehicles will affect the hub operating efficiency. Therefore, this research was based on the taxi picking up passengers in waiting area, explored the coupling relationship between passengers and vehicles, and realized the key simulation technology that one vehicle leaves with passengers then the vehicle behind it pulls into the empty parking space. This technology will increase the accuracy of hub simulation and improve 3D demonstration effect.

Keywords Transportation hub · Anylogic simulation · Relationship between passengers and vehicles

1 Introduction

As an important node of urban traffic network, integrated passenger transport hub is the place where various transportation modes and transportation lines converge. It has the characteristics of dense personnel, complex pedestrian flow direction and composition, complex organization and management influencing factors. The integrated passenger transport hub is a hub dominated by one transportation mode and coordinated by various transportation modes. For example, Beijing South Railway Station

J. Li · S. Hao (✉) · C. Jiao · N. Zhao · K. Wu
Research Institute of Highway, Ministry of Transport, 8 Xitucheng Rd., Haidian Dist., Beijing, China
e-mail: haosiyuan626@126.com

J. Li
e-mail: happyeveryday0903@126.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_39

523

is a large-scale passenger transport hub dominated by railway station, supplemented by subway and public transport, while Dongzhimen hub is a passenger transport hub dominated by subway and supplemented by public transport. With the rapid development and construction of the hubs, there are also a lot of urgent problems to be solved, such as the unreasonable transfer between various transport modes, resulting in the low service rate of the hub. Comprehensive passenger transport hub is the carrier of various modes of transportation, and all kinds of transportation modes serve passengers. Therefore, as the first service object of integrated passenger transport hub, the coupling relationship between passengers and transport modes affects the service efficiency of the whole hub. At present, simulation is often used in the design and operation stage of the hub to find the bottleneck and propose optimization measures. Most of the traffic simulation studies were vehicle traffic or pedestrian traffic separately, and the coupling simulation between passengers and vehicles was less. The accurate passenger vehicle coupling simulation can organically combine the vehicles and passengers in the hub, so as to improve the accuracy of the whole hub traffic simulation. Therefore, this paper focuses on the key technologies of coupling relationship between passengers and vehicles to improve the accuracy of traffic simulation in the hub.

The research on pedestrian simulation system began at the end of the twentieth century. Since then, the software specially used for pedestrian simulation has been gradually applied. At present, the pedestrian simulation software developed and applied abroad mainly includes Building, EXODUS, EGRESS, Legion, MYRIAD, PAXPORT, SIMULEX, VISSIM, Anylogic, etc. Still [1] carefully studied and considered the problems of pedestrian movement and safety in congested conditions, compared and analyzed the shortcomings and advantages of existing research results on pedestrian traffic, and used Legion for simulation analysis. Seti [2] carried out simulation analysis on the passenger flow in the hub, found the bottleneck position of the hub traffic operation through the simulation results, found out the influencing factors of the bottleneck and optimizes it through the analysis of the simulation parameters, so as to improve the passenger flow transfer efficiency of the hub. Gatersleben et al. [3], Lin and Trani [4], Seti and Hutchinson [5] Takakuwa and Oyama [6] conducted simulation research on passenger transfer in the hub by using simulation method, analyzed passenger transfer bottleneck, and proposed measures to eliminate bottleneck.

Jia [7] comprehensively sorted out and analyzed the relevant simulation modeling methods of integrated passenger transport hub, put forward the basic theory and method of simulation modeling, and combined with the actual case, carried out the simulation application. Li [8] conducted microscopic simulation on passenger flow of hub by using VISSIM software, so as to find out the bottleneck area of passenger transfer streamline. After giving optimization measures and simulating optimized streamline, the influencing factors of passenger transfer operation were summarized. Li [9] developed a new pedestrian simulation program. Based on the analysis of mainstream pedestrian simulation models, the relevant parameters are recalibrated and a new simulation program MTR-PedSIM is developed. Cui [10] used Direct3D program to conduct dynamic simulation of hub in actual cases. Based on the created

simulation environment, the technology of pedestrian height display and special area display was improved, and the two-dimensional simulation function was realized. Zhang [11] proposed measures for the design and optimization of passenger flow line in the hub through in-depth study on the passenger flow line, and applied Anylogic simulation software to study the passenger flow line in the hub.

Through the analysis of domestic and foreign scholars on pedestrian simulation research status, it can be seen that foreign pedestrian simulation research is far earlier than domestic. A lot of pedestrian simulation software has been developed in foreign countries, while the domestic research on pedestrian simulation is mainly improved on the basis of foreign simulation research.

For vehicle simulation research, in the theoretical discussion, system development, practical application and many other aspects of traffic simulation, foreign research started earlier and developed more mature than domestic. The study of traffic simulation abroad has roughly gone through three stages: 1960s, 1970s–1980s and 1990s. In the 1960s, the classic traffic simulation system is the TRANSYT macro traffic simulation model proposed by D. L. Robertson of Transportation and Road Research Laboratory. The representative work of this period in the 1970s and 1980s is the NETSIMU model, which was developed by the Federal Highway Administration. It has been widely used in the optimization of traffic planning schemes, traffic engineering and many other aspects. The simulation accuracy of the model has been greatly improved [12]. Since the 1990s, the concept of ITS (Intelligent Traffic System) has become a hot spot of scientific research in the field of traffic engineering. Many enterprises and scientific research institutions have carried out theoretical research, software development and practical application of traffic simulation system based on this concept. A large number of traffic simulation models and software have emerged. Up to now, TransModeler of the United States, VISSIM of Germany and PARAMICS of the UK are still widely used [13].

However, China began to devote itself to traffic simulation research in 1980s. Compared with foreign countries, there is still a gap in comprehensive research capacity, which needs continuous efforts to catch up. At present, the existing models mainly include TJTS (Tongji Traffic Simulation) developed by Tongji University and NITS (Network Integrated Traffic Simulation) developed by Northern Jiaotong University [14, 15].

On the basis of VISSIM traffic simulation software platform, Li et al. [16] selected reasonable indicators to conduct feasibility study and evaluation on the planning scheme of the study area, providing a good method for solving the operability study of urban road network planning scheme or traffic project. On the basis of TransModeler simulation system, Li [17] built a traffic simulation model of small and medium-sized cities in spatial scale, and combined with the field traffic survey data, made a detailed analysis on the causes of unbalanced traffic flow in the study area in different periods and directions.

According to the above research status at home and abroad, most of the simulation is only for internal pedestrians, or only for external vehicles, few researches on the coupling relationship between passengers and vehicles. Therefore, in the context of the rapid development of hub construction, aiming at the vacancy of the current

passengers and vehicle coupling simulation, taking taxis as an example, this paper focuses on the key technology of simulation of vehicle leaving with passengers and driving into empty parking space.

2 Key Technology of Coupling Relationship Between Passengers and Vehicles

The research was taking the coupling relationship between taxi vehicle and passenger as an example. The coupling relationship belongs to the mode of many to one, that is, the 1–3 passengers will choose the same taxi to get on at the same time, and the vehicle with passengers will leave immediately, and the taxis waiting for passengers will arrive at the empty parking space and continue to wait for passengers. There are three key technical problems in the coupling relationship between passengers and vehicles will be solved in this paper: 1–3 passengers get on one taxi at the same time; the taxi leaves immediately after carrying passengers; when the rear taxi judges that there is a free parking space, it arrives at the parking space and continues to wait for passengers.

3 Key Technology Implementation

3.1 Introduction of Anylogic Simulation Software

At present, the most commonly used pedestrian simulation softwares are Legion, Steps, SimWalk and AnyLogic. The first two are based on Cellular Automata Model and the latter two are based on Social Force Model. However, the first three softwares cannot be re-developed but Anylogic can. Because integrated passenger transport hub we studied was a complicated system, the secondary development function was needed, therefore, the Anylogic was chosen as the pedestrian simulation software in this study.

Anylogic is a multipurpose commercial simulation software created by XJ Technologies Company of Russia. It consists of basic simulation platform, pedestrian library, process library and so on. AnyLogic pedestrian library adopts flowchart method to create logical relation of simulation model. Besides, the flowchart can realize any correct logic, the logic relationship and the environment model can form a complete simulation model.

Anylogic pedestrian simulation is mainly realized through the functions in the pedestrian library. Different from traditional pedestrian simulation software, Anylogic does not need to give a complete path for pedestrians to walk. Instead, it simulates the model through Social Force Model. In Social Force Model, pedestrians have self-incentive mechanism, and the walking of each pedestrian is the

result of several factors. Such as the tendency of pedestrians to accelerate to their own comfortable speed, the tendency of pedestrians to keep a distance from other pedestrians or obstacles, and the tendency of pedestrians to approach specific pedestrians. While providing a Social Force Model as the basis for pedestrian simulation, Anylogic provides a high-degree of freedom development environment, which can achieve a high degree of customization, making the simulation process and results closer to the real situation.

Since Anylogic is a multi-purpose simulation software, it is not a professional pedestrian traffic simulation software. Pedestrians and vehicles are located in pedestrian library and road traffic library respectively. If they want to be coupled, they need to be redeveloped through programming.

3.2 Building Simulation Model

Take one floor of a large hub in China as an example. This floor is the transfer floor of intercity railway. The intercity railway once every 10 min. Passengers can take taxi on this floor to leave, or they can go to other floors to transfer buses or coaches. According to the design scheme, there are 12 taxi parking spaces, 6 in the South and 6 in the north. In order to compare and analyze the situation of considering coupling relationship and not considering coupling relationship, two taxi waiting areas with the same size and queuing mode are established, as shown in Fig. 1.

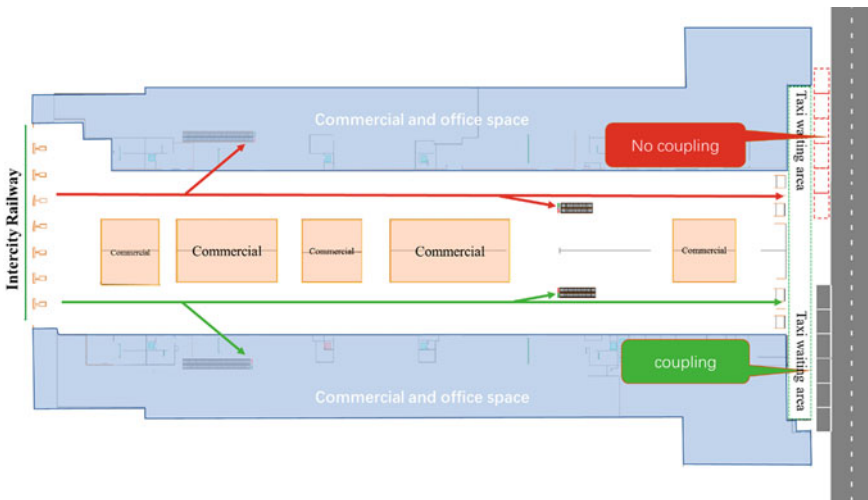


Fig. 1 Sketch map of simulation scene

3.3 Simulation Conditions

1. Assuming that the number of taxis is sufficient, the taxi parking space is in continuous turnover.
2. Coupling relationship: When the vehicle arrives at the parking space, 1–3 passengers leave by one taxi. The taxi waiting behind will drive into the parking space and continue to drive away with passengers. Repeat the process.
3. No coupling relationship: Taxi passengers disappear directly after passing through the taxi waiting area.

3.4 Simulation Logic Construction

The pedestrian and vehicle simulation flow charts are built in the main agent by using the pedestrian library and road traffic library of Anylogic (as shown in Fig. 2 and Fig. 3). The process of state diagram is established in car agent, and different action states such as vehicle arriving at parking space and taking passengers driving away are realized by programming. Based on the simulation condition of 3.3, the coupling relationship between passengers and vehicles is realized as shown in Fig. 4. The case without coupling relationship is shown in Fig. 5.



Fig. 2 Pedestrian simulation flowchart

Fig. 3 Vehicle simulation flowchart

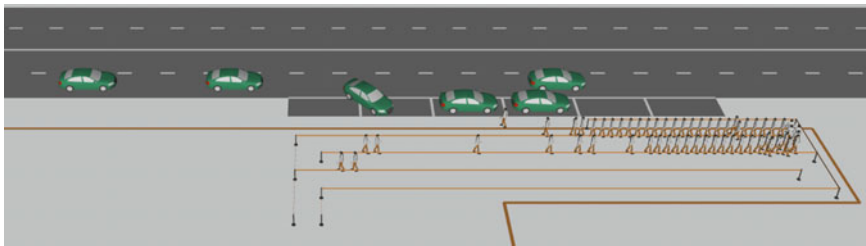
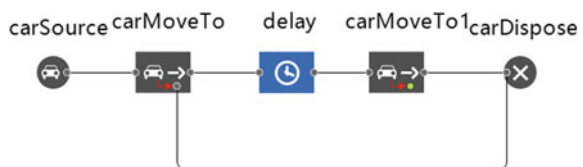


Fig. 4 Taxi waiting area (coupling relationship)

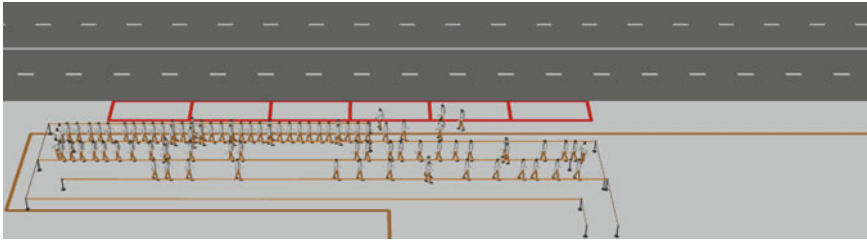


Fig. 5 Taxi waiting area (no coupling relationship)

4 Simulation Results Analysis

Test the difference between the two cases by traffic loading. This simulation tested five situations: the passengers who get off each intercity railway to transfer to taxi are 50, 100, 150, 200 and 250 respectively. Take the 50, 150, 250 passengers as examples, the simulation results are as follows:

When 50 passengers get off each train and transfer to a taxi, the density distribution diagrams without coupling relationship and with coupling relationship are shown in Fig. 6. In both cases, when passengers from the next train arrive at the taxi waiting area, all the taxi passengers on the previous train have dissipated. And the maximum waiting area density without coupling relationship is slightly higher than the case with coupling relationship.

When 150 passengers get off each train and transfer to taxi, the density distribution diagrams without coupling and with coupling are shown in Fig. 7. In both cases, when the passengers from the last train arrive at the taxi waiting area, the waiting passengers of the previous train have not dissipated completely, so there will be no density of zero. But the obvious peaks and troughs can be explained that it has a dissipation process in taxi waiting area. The maximum waiting area density with coupling relationship is higher than without coupling relationship. That is the service capacity with coupling relationship is lower than that without coupling relationship.

When 250 passengers get off each train and transfer to taxi, the density distribution diagrams without coupling and with coupling are shown in Fig. 8. There is no obvious

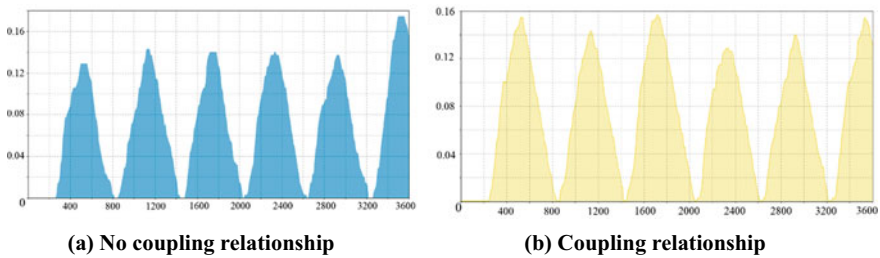


Fig. 6 Density distribution diagrams with 50 passengers

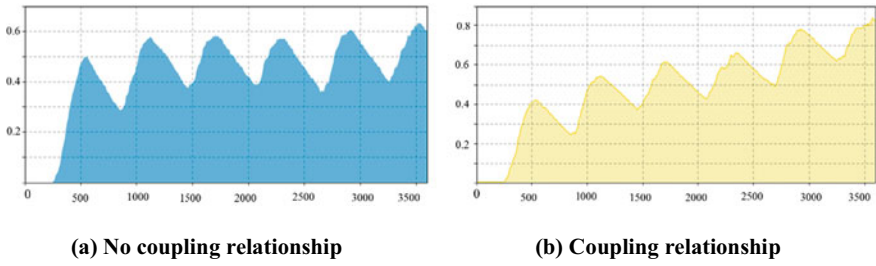


Fig. 7 Density distribution diagrams with 150 passengers

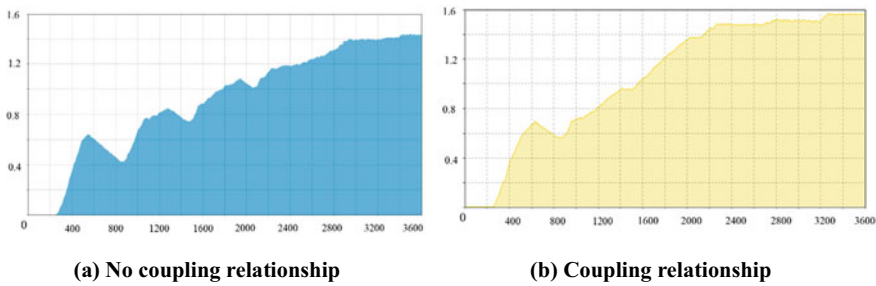


Fig. 8 Density distribution diagrams with 250 passengers

peaks and troughs, but the density keeps increasing. The maximum waiting area density basically the same in both cases. But the average density with coupling relationship is higher than without coupling relationship.

Counted the arrival and departure of passengers within one hour under the condition of different taxi passengers. The statistical results are shown in Table 1.

When the number of taxi passengers is small, the service efficiency with coupling relationship is higher than that without coupling relationship. Because the case with coupling relationship takes into account the situation of 2–3 passengers sharing,

Table 1 One hour statistics of different conditions

No. of taxi passengers per 10 min	No coupling relationship			Coupling relationship		
	Arrive	Depart	Depart-to-arrive ratio (%)	Arrive	Depart	Depart-to-arrive ratio (%)
50	302	246	81	308	262	85
100	636	544	86	578	473	82
150	871	708	81	880	590	67
200	1170	699	60	1206	590	49
250	1472	701	48	1494	585	39

and when the number of passengers is small, taxi waiting passengers, so the queue passengers dissipate quickly.

As the number of passengers increases, the service efficiency with coupling relationship is lower than that without coupling relationship. Because there is a turnover time between the departure of the former taxi and the entry of the next taxi into the parking space, the case with coupling relationship becomes passengers waiting taxi. Without coupling relationship, the taxi turnaround time is not involved, so the queue passengers dissipate quickly.

When the number of taxi passengers is small, the departure rate of the two cases is basically the same, that is, whether considering the coupling relationship has little effect. When reaching their respective service limits, the service limit is about 700 passengers without considering the coupling relationship, and about 590 passengers when considering the coupling relationship.

5 Discussion and Conclusion

According to the results of Anylogic simulation, there is a big difference between considering the coupling relationship and not considering the coupling relationship. Considering the coupling is closer to the real situation, because it considers the situation of multiple passengers sharing and the time of vehicle turnover. It not only improves the simulation accuracy, but also makes the 3D simulation more realistic. It has guiding significance for future research on hub transfer efficiency and parking space planning and design.

However, this paper still needs to be improved. For example, the case mentioned in this paper does not consider the coupling relationship, assuming that passengers disappear directly after passing through the taxi waiting area. In practical application, different modelers will put forward different assumptions, and the final simulation results will be different.

In this paper, the key technologies of coupling between passenger and vehicle are basically realized, but there are still some spaces to be improved, such as the time for passengers to leave luggage is not considered, which will be improved in the future research.

References

1. Still GK (2000) Crowd dynamics. University of Warwick, Coventry
2. Seti JR, Hutehi BG (1994) Passenger-terminal simulation model. *J Transp Eng* 120(4):515–536
3. Gatersleben MR, van der Weij, Simon W (1999) Analysis and simulation of passenger flows in an airport terminal. In: Winter simulation conference proceedings. Phoenix, AZ, USA, 1999. Piscataway, NJ, USA, 1999, pp 1226–1231
4. Lin Y-D, Trani AA (2000) Airport automated people mover systems: analysis with a hybrid computer simulation model. *Transp Res Rec* 2000(1703):45–57

5. Seti JR, Hutchinson BG (1994) Passenger-terminal simulation model. *J Transp Eng* 120(4):517–535
6. Takakuwa S, Oyama T (2003) Simulation analysis of international-departure passenger flows in an airport terminal. In: Winter simulation conference proceedings. New Orleans, LA, USA, IEEE, pp 1627–1634
7. Jia H (2011) Key theory and method of simulation modeling for integrated transportation passenger hub. Science Press, Beijing
8. Li J (2011) Comprehensive transportation hubs dynamic evaluation method for passenger organization. Beijing Jiaotong University
9. Li D, Han B, Li H (2009) Microscope simulation on large scale railway station of passenger mustering and evacuation. *China Railway Sci* (03):119–120
10. Cui C, Chen B, Chen Z, Qin S (2010) Three-dimensional simulation of integrated transport hub for passenger traffic system. *J Transport Inf Saf* (02):8–12
11. Zhang S (2010) Research on model and method for passenger flow optimization in a large integrated traffic terminal. Beijing Jiaotong University
12. Rathi AK (1990) Urban network traffic simulation TRAFNETSIM program. *J Transp Eng* 116(6):734–743
13. Weng Y (2002) Modeling and simulation of urban mixed traffic system. Zhejiang University
14. Zhou Z (2000) Research on urban road traffic simulation. Tongji University
15. Chen Q, Xie A, Hu S (2001) Large scale urban traffic simulation system. *J Syst Simul* 06:831–833
16. Li D, He Y, Gao Y (2016) Evaluation of transportation planning schemes based on Vissim simulation. *Road Traffic Saf* 03:7–12+22
17. Li G (2013) Modeling and simulation of traffic control systems in small and Medium-sized cities based on transmodeler. Zhejiang University

Coevolution Dynamics Model of Urban Multimodal Transportation System Under Limited Exploitable Land Resources



Rong Li, Weili Kong, Liwen Zhang, and Peng Shu

Abstract Urban multimodal transportation system evolution is a complex dynamical process under limited exploitable land resources. In this paper, a five-level coevolution dynamics model among land development, population, roads, private vehicle and bus ownership is developed to capture the growth mechanisms of urban traffic systems. It is convergent under certain conditions through stability analysis. This model can capture the observed co-evolution law and predict the evolution trend of urban multimodal transportation system in the future by taking Beijing city as case study. It is of great significance to urban traffic planning and management.

Keywords Urban multimodal transportation · Coevolution dynamics model · Limited exploitable land resources

1 Introduction

With the acceleration of urbanization, the scale of urban development in developing countries, such as China, is expanding, the urban population is growing rapidly, and the traffic congestion is becoming more and more serious. Limited land resource is the constraint of the urban development. Population growth accelerates the evolution of urban road traffic system. Roads are the sources for the urban traffic structure (mainly contains private vehicles and public buses). The contradiction between traffic supply and demand is the root cause of traffic congestion. The experience of many developed

R. Li (✉)

School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
e-mail: 16114196@bjtu.edu.cn

R. Li · L. Zhang

Department of Traffic Management, Jiangxi Police Institute, Nanchang 330100, China

W. Kong

Nanchang City Passenger Transportation Management Office, Nanchang 330009, China

P. Shu

Shenzhen Institute of Technology, Shenzhen 518045, China

cities shows that blindly increasing road traffic supply can not solve the problem of traffic congestion, and developing multimodal transportation system is general trend.

For single modal traffic system evolution, Yamins et al. built a road network growing dynamics model that can generate typical belt-ways and star patterns in urban traffic network [1]. A co-evolution model of land use and transportation network is built as a bottom-up process [2]. Barthélemy et al. studied the interaction between population density evolution and the road network growth [3]. Li et al. built a co-evolution model between land use and traffic network design [4]. A city expansion model is proposed to reveal the coevolution process between population diffusion and road growth by Wu et al. [5]. Zhao et al. proposed a coevolution model by taking into account of the interplay between population distribution and road network [6]. A two-level logistic model is proposed to analyze the co-evolution of road networks and vehicle ownership [7]. Wu et al. proposed a coevolution model to consider the interaction among traveler, vehicle and roads [8]. Li et al. built a four-dimensional co-evolution model to reveal the growth mechanisms of urban traffic systems [9]. Zhao et al. studied the co-development of population distribution and road networks by considering the RNG and Fermat–Weber location problem as the connection mechanism [10]. Zhao et al. built a road network growth model with the consideration of population density distribution and CBD attraction [11]. Ding et al. showed that network growth is dominated by two key elementary elements: expansion and densification [12].

As we all know, it is very difficult to solve the problem of road traffic congestion in large cities only by considering single modal traffic infrastructure. It is imperative to build a urban multi-modal traffic network with reasonable structure, convenient access and efficient connection. In recent years, a large number of scholars have carried out research on multimodal transportation system. Nes investigated the consequences of multimodal travelling for designing multimodal transport networks [13]. Wan et al. developed a systematic methodology for multi-modal network design by considering traffic congestion and modal transfers [14]. Gonzales et al. addresses how street space should be allocated and how transport modes should be operated for different city structures by considering several costs: time, money, space, and externalities [15]. Zhang et al. investigated the multimodal network design problem that optimizes the private car network expansion and bus network design [16]. Zhang et al. examined the time-dependent bus scheduling problem in a multimodal transportation network environment [17].

The above researches on multimodal transportation system are mainly based on network optimization. In fact, urban traffic system is composed of human, vehicle, road and environment. Urban multimodal transportation system is a complex, giant and dynamical system, co-evolved with city population and land development. Roads are the supply of the urban traffic system, while the traffic structure which includes cars, buses and taxis is the demand. In this paper, the factors considered are cars and buses which are the main modes of road transportation. Urban transportation system evolution with other urban aspects as land development, population and so on is a hot issue in the current traffic system science research.

As for the evolution of urban multimodal traffic system, the previous researches were explored from two different fields: the one is traffic network evolution based on complex network science, the other is about traffic system (always contains several statistical indicators) evolution based on system dynamics. With the development of multi-layer complex network theory [18–20], multi-layer network-based traffic system evolution has attracted many attentions [21]. Ding et al. proposed a multi-layer urban network model with the interaction between lower layer and upper layer network [22]. Ding et al. studied the network evolution trends and influences of urban traffic dynamics based on multi-layer network theory [23]. Ding et al. performed a simulative experiment and the optimized multilayer urban traffic network structure under different contexts [24]. As for the system dynamics model, Xu et al. built a coevolution dynamics model to simulate the coevolution mechanisms between road network, private and public traffic structure [25].

Although there have been numerous contributions on the multimodal traffic system evolution, there have been also many shortages as follow. First, most of the multimodal traffic system evolution is about network evolution, while there are very few of system dynamics. Secondly, Xu et al. built a coevolution dynamics model to simulate the coevolution mechanisms between road network, private and public traffic structure [25], but not considered the influence of land development [9]. In this paper, we develop a five-dimensional co-evolution dynamics model among land use, population, roads and private car and bus ownership based on two previous researches by Li et al. [9] and Wu et al. [25]

The next section shows the co-evolution dynamics model. In Sect. 3, we analyze the stability of this model. A case study is exhibited in Sect. 4. The conclusion section reveals a summary of the research results and perspectives.

2 Model

The main model framework continue to use the system dynamics model built by Li et al. [9], meanwhile combine the traffic structure (contains private vehicles and public buses) coevolution model structure as showed in Sect. 2.2.

2.1 Notations

The notations used in this paper are described in Table 1.

Table 1 The parameters and their descriptions in this model

Parameters	Descriptions	Parameters	Descriptions
\bar{L}	The land area in the city	γ_4	The growth coefficient of private vehicle ownership
L	The built-up area in the city	γ_5	The growth coefficient of public bus ownership
P	The population in the city	α_1	The maximum ratio of land development
R	The road area in the city	α_2	The maximum population density
V	The private vehicle ownership in the city	α_3	The maximum ratio of road in the built-up area
B	The public bus ownership in the city	α_4	The maximum carrying capacity of vehicles per unit road area
γ_1	The growth coefficient of the built-up area	α_5	The maximum carrying capacity of bus per unit road area
γ_2	The growth coefficient of population	σ_1	Feedback parameter of private vehicle ownership to public bus ownership
γ_3	The growth coefficient of the road area	σ_2	Feedback parameter of public bus ownership to private vehicle ownership

2.2 Co-evolution Model

In this model, we consider the evolution of land development with an upper limit to exploitable land area in the city. The population and road area evolution is driven by land development evolution. The private vehicle and public bus ownership evolution is influenced by population and each other. Land development, population and road area evolution dynamics can be expressed as Eqs. (1)–(3), which is same as the model built by Li et al. [9].

$$\frac{dL}{dt} = \gamma_1 L \left(1 - \frac{L}{\alpha_1 \bar{L}} \right) \tag{1}$$

$$\frac{dP}{dt} = \gamma_2 P \left(1 - \frac{P}{\alpha_2 L} \right) \tag{2}$$

$$\frac{dR}{dt} = \gamma_3 R \left(1 - \frac{R}{\alpha_3 L} \right) \tag{3}$$

For the private vehicle and public bus ownership evolution dynamics, this paper combine the model built by Li et al. [9] and Wu et al. [25], which consider that population growth promotes private motor vehicle and public bus ownership evolution, the traffic (private vehicle and public bus) carrying capacity is linear to the road area

and the interaction between private motor vehicle and public bus ownership evolution by logistic equation.

$$\frac{dV}{dt} = \gamma_4 P V \left(1 - \frac{V}{\alpha_4 R} + \sigma_1 \frac{B}{\alpha_5 R} \right) \tag{4}$$

$$\frac{dB}{dt} = \gamma_5 P B \left(1 - \frac{B}{\alpha_5 R} + \sigma_2 \frac{V}{\alpha_4 R} \right) \tag{5}$$

Thus, the general coevolution dynamic model among land development, population, road, private vehicle and public bus can be expressed as Eq. (6):

$$\begin{cases} \frac{dL}{dt} = \gamma_1 L \left(1 - \frac{L}{\alpha_1 \bar{L}} \right) \\ \frac{dP}{dt} = \gamma_2 P \left(1 - \frac{P}{\alpha_2 \bar{L}} \right) \\ \frac{dR}{dt} = \gamma_3 R \left(1 - \frac{R}{\alpha_3 \bar{L}} \right) \\ \frac{dV}{dt} = \gamma_4 P V \left(1 - \frac{V}{\alpha_4 R} + \sigma_1 \frac{B}{\alpha_5 R} \right) \\ \frac{dB}{dt} = \gamma_5 P B \left(1 - \frac{B}{\alpha_5 R} + \sigma_2 \frac{V}{\alpha_4 R} \right) \\ \gamma_i \geq 0, \alpha_i \geq 0, \sigma_i \geq 0 \end{cases} \tag{6}$$

3 Stability Analysis

Set $\frac{dL}{dt} = 0, \frac{dP}{dt} = 0, \frac{dR}{dt} = 0$ and $\frac{dV}{dt} = 0, \frac{dB}{dt} = 0$, then the equilibrium solutions can be obtained. There are 32 equilibrium points, but we only consider the untrivial solution set forth below:

$$E = \left(\alpha_1 \bar{L}, \alpha_1 \alpha_2 \bar{L}, \alpha_1 \alpha_3 \bar{L}, \frac{1 + \sigma_1}{1 - \sigma_1 \sigma_2} \alpha_1 \alpha_3 \alpha_4 \bar{L}, \frac{1 + \sigma_2}{1 - \sigma_1 \sigma_2} \alpha_1 \alpha_3 \alpha_5 \bar{L} \right) \tag{7}$$

We transform the variables $L' = L - \alpha_1 \bar{L}, P' = P - \alpha_1 \alpha_2 \bar{L}, R' = R - \alpha_1 \alpha_3 \bar{L}, V' = V - \frac{1 + \sigma_1}{1 - \sigma_1 \sigma_2} \alpha_1 \alpha_3 \alpha_4 \bar{L}$ and $B' = B - \frac{1 + \sigma_2}{1 - \sigma_1 \sigma_2} \alpha_1 \alpha_3 \alpha_5 \bar{L}$; thus,

$$\left\{ \begin{aligned} \frac{dL'}{dt} &= \gamma_1(L' + \alpha_1\bar{L})\left(1 - \frac{L'+\alpha_1\bar{L}}{\alpha_1\bar{L}}\right) \\ \frac{dP'}{dt} &= \gamma_2(P' + \alpha_1\alpha_2\bar{L})\left(1 - \frac{P'+\alpha_1\alpha_2\bar{L}}{\alpha_2(L'+\alpha_1\bar{L})}\right) \\ \frac{dR'}{dt} &= \gamma_3(R' + \alpha_1\alpha_3\bar{L})\left[1 - \frac{R'+\alpha_1\alpha_3\bar{L}}{\alpha_3(L'+\alpha_1\bar{L})}\right] \\ \frac{dV'}{dt} &= \gamma_4(P' + \alpha_1\alpha_2\bar{L})(V' + \frac{1+\sigma_1}{1-\sigma_1\sigma_2}\alpha_1\alpha_3\alpha_4\bar{L})\left[1 - \frac{V'+\frac{1+\sigma_1}{1-\sigma_1\sigma_2}\alpha_1\alpha_3\alpha_4\bar{L}}{\alpha_4(R'+\alpha_1\alpha_3\bar{L})} + \sigma_1\frac{B'+\frac{1+\sigma_2}{1-\sigma_1\sigma_2}\alpha_1\alpha_3\alpha_5\bar{L}}{\alpha_5(R'+\alpha_1\alpha_3\bar{L})}\right] \\ \frac{dB'}{dt} &= \gamma_5(P' + \alpha_1\alpha_2\bar{L})(B' + \frac{1+\sigma_2}{1-\sigma_1\sigma_2}\alpha_1\alpha_3\alpha_5\bar{L})\left[1 - \frac{B'+\frac{1+\sigma_2}{1-\sigma_1\sigma_2}\alpha_1\alpha_3\alpha_5\bar{L}}{\alpha_5(R'+\alpha_1\alpha_3\bar{L})} + \sigma_2\frac{V'+\frac{1+\sigma_1}{1-\sigma_1\sigma_2}\alpha_1\alpha_3\alpha_4\bar{L}}{\alpha_4(R'+\alpha_1\alpha_3\bar{L})}\right] \\ 0 \leq \gamma_i \leq 1, \alpha_i \geq 0, \sigma_i \geq 0 \end{aligned} \right. \tag{8}$$

On the other hand, there are also 32 equilibrium points, and $E' = (0, 0, 0, 0, 0)$ corresponds to E for Eq. (6). Thus, the stability of E for Eq. (6) is equivalent to E' for Eq. (8). The Jacobian matrix of Eq. (8) is as follows:

$$A = \begin{bmatrix} -\gamma_1 & 0 & 0 & 0 & 0 \\ \gamma_2\alpha_2 & -\gamma_2 & 0 & 0 & 0 \\ \gamma_3\alpha_3 & 0 & -\gamma_3 & 0 & 0 \\ 0 & 0 & \gamma_4\alpha_1\alpha_2\alpha_4\bar{L}\frac{1+\sigma_1}{1-\sigma_1\sigma_2} & -\gamma_4\alpha_1\alpha_2\bar{L}\frac{1+\sigma_1}{1-\sigma_1\sigma_2} & \gamma_4\alpha_1\alpha_2\bar{L}\frac{\sigma_1+\sigma_1^2}{\alpha_5(1-\sigma_1\sigma_2)} \\ 0 & 0 & \gamma_5\alpha_1\alpha_2\alpha_5\bar{L}\frac{1+\sigma_2}{1-\sigma_1\sigma_2} & \gamma_5\alpha_1\alpha_2\alpha_5\bar{L}\frac{\sigma_2+\sigma_2^2}{\alpha_4(1-\sigma_1\sigma_2)} & -\gamma_5\alpha_1\alpha_2\bar{L}\frac{1+\sigma_2}{1-\sigma_1\sigma_2} \end{bmatrix} \tag{9}$$

The characteristic equation of A is $|\lambda I - A| = 0$. Its solutions are $\lambda_1 = -\gamma_1$, $\lambda_2 = -\gamma_2$, $\lambda_3 = -\gamma_3$, $\lambda_4 = \frac{a+\alpha_1\alpha_2\bar{L}\sqrt{b}}{2(\sigma_1\sigma_2-1)}$, $\lambda_5 = \frac{a-\alpha_1\alpha_2\bar{L}\sqrt{b}}{2(\sigma_1\sigma_2-1)}$. Where

$$a = \gamma_4\alpha_1\alpha_2\bar{L} + \gamma_5\alpha_1\alpha_2\bar{L} + \gamma_4\alpha_1\alpha_2\sigma_1\bar{L} + \gamma_5\alpha_1\alpha_2\sigma_2\bar{L} \tag{10}$$

$$\begin{aligned} b &= \gamma_4^2\sigma_1^2 + 2\gamma_4^2\sigma_1 + \gamma_4^2 + 4\gamma_4\gamma_5\sigma_1^2\sigma_2^2 + 4\gamma_4\gamma_5\sigma_1^2\sigma_2 \\ &\quad + 4\gamma_4\gamma_5\sigma_1\sigma_2^2 + 2\gamma_4\gamma_5\sigma_1\sigma_2 - 2\gamma_4\gamma_5\sigma_2 - 2\gamma_4\gamma_5 \\ &\quad + \gamma_5^2\sigma_2^2 + 2\gamma_5^2\sigma_2 + \gamma_5^2 \end{aligned} \tag{11}$$

According to Li et al., the zero solution is asymptotically stable if all the roots of the characteristic equation have a negative real part [9]. For this characteristic equation, its solutions all have negative real part when

$$\gamma_1 > 0, \quad \gamma_2 > 0, \quad \gamma_3 > 0, \quad b > 0, \quad \frac{a}{\sigma_1\sigma_2 - 1} < 0 \tag{12}$$

Table 2 The range of values and calibrated parameter values of the parameters

Parameters name	Value range	Calibrated parameter value	Parameters name	Value range	Calibrated parameter value
γ_1	$(0, +\infty)$	0.1054	α_1	$(0, 0.1)$	0.1000
γ_2	$(0, +\infty)$	0.1631	α_2	$(0, 2)$	2.0000
γ_3	$(0, +\infty)$	0.1625	α_3	$(0, 1)$	0.0989
γ_4	$(0, +\infty)$	0.0001	α_4	$(0, 5.6)$	4.4789
γ_5	$(0, +\infty)$	2.9948	α_5	$(0, 2.8)$	0.0179
σ_1	$(0, 1)$	0.4445	σ_2	$(0, 1)$	0.5719

4 Case Study

4.1 Parameter Calibration

The territorial area of Beijing city $\bar{L}_{Beijing} = 16410.54$, the parameters involved in the co-development model is calibrated by the evolution data of built-up area, population, road area, and private car and public bus ownership (from 1988 to 2012) through the least squares principle. The range of values and calibrated parameter values of the parameters in the cities are shown in Table 2.

4.2 Simulation Results

The parameters value in this model can be calibrated by real data from 1988 to 2012. The evolutionary processes from 2013 to 2017 can be predicted by the calibrated model. The comparison of the results and the real data in these years is exhibited in Table 3.

The comparison of the simulation and real evolution of the built-up area, population, road area and private car and public bus ownership in Beijing are showed in Fig. 1. With respect to the built-up area, the absolute of relative tolerance is less than 3%; for population, the road area, private car and public bus ownership, the absolute is less than 17%, 12%, 41% and 36%. Therefore, the coevolution model can effectively predict the co-development process in Beijing.

5 Conclusion and Perspectives

In this paper, a five-level coevolution model is proposed to capture the interaction characteristic among multimodal transportation system under limited exploitable

Table 3 Comparative analysis results of the prediction and the real data

Year	Built up area			Population		
	Real data	Prediction result	Relative tolerance (%)	Real data	Prediction result	Relative tolerance (%)
2013	1306.50	1335.01	2.18	21.15	21.45	1.41
2014	1385.60	1360.39	-1.82	21.52	22.13	2.85
2015	1401.00	1384.07	-1.21	21.71	22.79	5.02
2016	1420.00	1406.10	-0.98	21.73	23.44	7.88
2017	1450.00	1426.53	-1.62	20.71	24.07	16.22
Year	Road area			Private car ownership		
	Real data	Prediction result	Relative tolerance (%)	Real data	Prediction result	Relative tolerance (%)
2013	96.11	101.29	5.39	5.44	6.07	11.70
2014	100.02	105.08	5.06	5.59	6.66	19.05
2015	100.29	108.77	8.45	5.62	7.23	28.61
2016	102.75	112.34	9.34	5.72	7.78	36.02
2017	103.47	115.79	11.91	5.91	8.30	40.42
Year	Public bus ownership					
	Real data	Prediction result	Relative tolerance (%)			
2013	27.59	31.99	15.96			
2014	28.33	34.02	20.08			
2015	28.31	36.00	27.14			
2016	27.89	37.86	35.75			
2017	30.97	39.67	28.11			

Note The real data for the built-up area in 2010 are replaced by the mean value of the years 2009 and 2011, which is not available in the statistical data

land resources in the city by combining the model in Li et al. [9] and Wu et al. [25]. For the model, on the one hand, we consider the urban multimodal transportation system (population, road area, private vehicle and public bus ownership) evolution driven by the land development evolution under upper limit to a given ratio of urban land area. On the other hand, this model is asymptotically stable and converges to a fixed point under certain conditions through proof. Taking Beijing city as case study, the simulation results reveal that the model can effectively capture the coevolution trajectory of land development, population, and road area, private car and public bus ownership and predict the evolution trends of multimodal transportation system.

Despite all this, there are also some idealized assumptions in this model. For example, the carrying capacities of population and traffic are assumed to be linear to the built-up area and road area, which may be nonlinear in the real situation. The given exploitable upper limit to land area may change with the development of science and technology. Recently, the microscopic multimodal traffic flow has been a

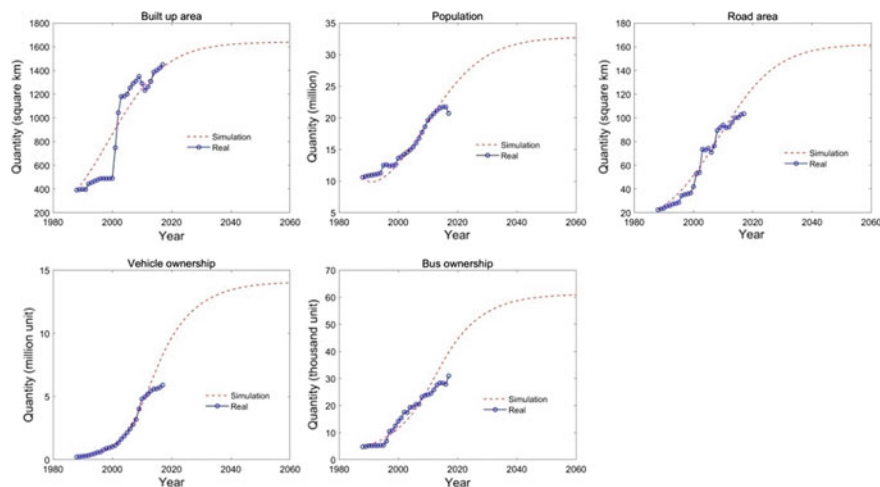


Fig. 1 Co-evolution of the built-up area, population, road area, private car and public bus ownership in Beijing city

hot topic, how to combine the macroscopic evolution with the microscopic behavior is a challenging work. So there are also many complex problems to be solved.

Acknowledgements This work is supported by the Youth fund of Jiangxi Social Science Planning Project (19GL46), Key projects of science and technology research project of Jiangxi Provincial Department of Education (GJJ191027), Youth fund of Humanities and Social Science Research Project in Colleges and Universities of Jiangxi Province (GL20216) and Youth fund of science and technology research project of Jiangxi Police Institute (2018QN005).

Compliance with Ethical Standards Conflict of Interest: The authors declare that there is no conflict of interests regarding the publication of this article.

References

1. Yamins D, Rasmussen S, Fogel D (2003) Growing urban roads. *Networks Spatial Econ* 3:69–85
2. Levinson D, Xie F, Zhu SJ (2007) The co-evolution of land use and road networks. *Transp Traffic Theory* 17:839–859
3. Barthélemy M, Flammini A (2009) Co-evolution of density and topology in a simple model of city formation. *Netw Spat Econ* 9:401–425
4. Li TF, Wu JJ, Sun HJ, Gao ZY (2016) Integrated co-evolution model of land use and traffic network design. *Netw Spat Econ* 16:579–603
5. Wu JJ, Li R, Ding R, Li TF, Sun HJ (2017) City expansion model based on population diffusion and road growth. *Appl Math Model* 43:1–14
6. Zhao FX, Wu JJ, Sun HJ, Gao ZY, Liu RH (2016) Population-driven urban road evolution dynamic model. *Netw Spat Econ* 16:997–1018
7. Xu M, Ye Z, Shan X (2016) Modeling, analysis, and simulation of the co-development of road networks and vehicle ownership. *Phys A: Stat Mech Appl* 442:417–428

8. Wu JJ, Xu M, Gao ZY (2014) Modeling the coevolution of road expansion and urban traffic growth. *Adv Complex Syst* 17:1450005-1–1450005-18
9. Li R, Wu J, Liu H et al (2019) Crowded urban traffic: co-evolution among land development, population, roads and vehicle ownership. *Nonlinear Dyn* 95(4):2783–2795
10. Zhao FX, Zhao FW, Sun H (2019) A coevolution model of population distribution and road networks. <https://doi.org/10.1016/j.physa.2019.04.096>
11. Zhao F, Sun H, Wu J, Gao Z, Liu R (2016) Analysis of road network pattern considering population distribution and central business district. *PLoS ONE* 11(3):e0151676. <https://doi.org/10.1371/journal.pone.0151676>
12. Ding J, Qin R, Guo N et al (2019) Urban road network growth model based on RNG proximity graph and angle restriction. *Nonlinear Dyn* 96:2281–2292
13. Van Nes R (2002) Design of multimodal transport networks, a hierarchical approach. TRAIL Thesis Series T2002/5, DUP, Delft University, The Netherlands (2002)
14. Wan QK, Hong KL (2009) Congested multimodal transit network design. *Public Transp* 1:233–251
15. Gonzales E (2011) Allocation of space and the costs of multimodal transport in cities. Ph.D. thesis, University of California, Berkeley
16. Zhang L, Yang H, Wu D, Wang D (2014) Solving a discrete multimodal transportation network design problem. *Transp Res Part C: Emerg Technol* 49:73–86
17. Zhang F, Liu W (2020) Responsive bus dispatching strategy in a multi-modal and multi-directional transportation system: a doubly dynamical approach. *Transp Res Procedia* 113:21–37
18. Nicosia V, Bianconi G, Latora V, Barthélemy M (2013) Growing multiplex networks. *Phys Rev Lett* 111:058701
19. Kivelä M et al (2014) Multilayer networks. *J Complex Netw* 2(3):203–271
20. Boccaletti S et al (2014) The structure and dynamics of multilayer networks. *Phys Rep* 544(1):1–122
21. Ding R, Ujang N, Hamid HB et al (2019) Application of complex networks theory in urban traffic network researches. *Netw Spat Econ* 19:1281–1317
22. Ding R, Ujang N, Hamid HB, Manan MSA, Li R, Wu JJ (2017) Heuristic urban transportation network design method, a multilayer coevolution approach. *Phys A Stat Mech Appl* 479:71–83
23. Ding R, Ujang N, Hamid HB, Manan MSA, He Y, Li R, Wu JJ (2018) Detecting the urban traffic network structure dynamics through the growth and analysis of multi-layer networks. *Phys A Stat Mech Appl* 503:800–817
24. Ding R, Yin J, Dai P, Jiao L, Li R, Li T, Wu J (2019) Optimal topology of multilayer urban traffic networks. *Complexity* 2019, Article ID 4230981, 19 p
25. Wu JJ, Xu M, Gao ZY (2013) Coevolution dynamics model of road surface and urban traffic structure. *Nonlinear Dyn* 73:1327–1334

Traffic Congestion Improvement Scheme of the Corridor Between Beijing Deputy Center and Central Urban Area



Chuanjiao Sun and Panyi Wei

Abstract In order to solve the congestion problem between Beijing deputy Center and Central Urban Area after the movement of Beijing government, the main corridors between the two areas are studied on the main congestion index and the importance of the area. A model about area affections are constructed by the congestion and the importance. Based on the model, the main congestion locations are given and the improvement measures are listed.

Keywords Traffic congestion · Deputy center · Central urban area

1 Introduction

According to the new edition of Beijing's urban plan, Tongzhou, Beijing's deputy center, is an important carrier of non-capital functions. Together with Xiong'an new area, it will form "two wings" of Beijing in the future and will become a model of urban construction and development in the new era. The Central Committee of the Communist Party of China and the State Council announced in the "Regulatory Detailed Plan for Beijing's Deputy Urban Center (at the Block Level) (2016–2035)" that by 2035, the city will initially be built into a harmonious and livable modern city with core competitiveness, cultural charm, urban vitality and world-class features. After the relocation of the Beijing municipal government, it is worth carrying out specific analysis and research on the impact of the existing transportation corridor to find the possible congestions for further consideration.

C. Sun (✉) · P. Wei
Road Safety Research Center, RIOH, Beijing 100088, China

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_41

543

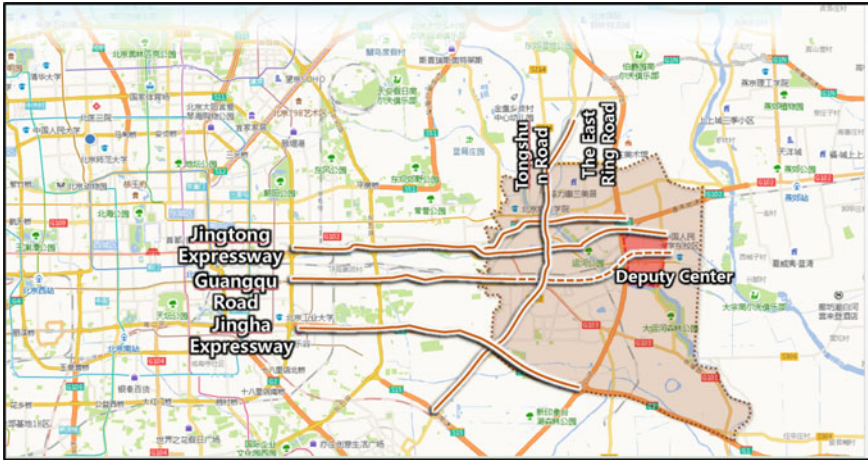


Fig. 1 The main corridor between Beijing deputy center and central urban area

2 Main Transport Corridor

As it can be seen from Fig. 1, the main transport corridors between Beijing deputy center and central urban area include: Chaoyang north road, Chaoyang road, Jingtong expressway, Guangqu road. These corridors are all the east-west directions. There are 2 north-south roads, which are Tongshun road and East 6th ring road.

By analysis on the past one year Gaode travel data, the congestion indexes are given for further study. The highest congestion index 3.11 is Sanhui interchange, which intersects with East Sixth Ring Road with an important control effect restricting Jingtong entering the core area.

The second congestion index 1.5 appears in Xinhua West Street at intersections with Xiguan Second Lane and Xincang Road.

All the other road sections are relatively smooth with congestion index ranging from 1 to 1.5.

3 Main Congestion Road

See Table 1.

Table 1 The top 20 congestion road

Road name	Congestion index	Length (m)
Ma tuan Road	2.72	320
Chezhan Road	2.71	1092
Wuzhi xueyuan Road	2.59	634
Yongchang south road	2.58	565
Maqiao north street	2.56	1973
Xinhua west street	2.55	2420
Matuan Road west part	2.38	320
West ring south road	2.35	1841
Luyuan south street	2.29	4038
Xinhua south road	2.25	1836
Tonghua road	2.24	2466
Xin hua west street auxiliary lane	2.24	545
Yongshun south street	2.24	1283
Chaoyang north street	2.19	2752
Chezhan Road (near Xiying street)	2.19	1092
Yongchang south road	2.14	550
6 south road	2.10	518
Tonghui south road	2.09	978
Guangqu road	2.08	353
Siyuanting street	2.07	377

4 Analysis of Road Congestion Impact Index Based on Location Importance

In order to further analyze the impact of different road congestion on commuter travel in urban sub-centers and administrative office areas, as well as the impact on non-commuter travel in urban sub-centers, a road congestion impact index analysis model for urban sub-centers based on location importance is established by considering factors such as road location, road connectivity, road administrative grade, congestion index, etc.

$$TII = T_x (af_{(x1)} + bf_{(x2)} + cf_{(x3)})$$

TII is the congestions impact index; T_x is the congestions index. $f_{(x1)}$ is the road connectivity index; $f_{(x2)}$ is the road grade impact index; $f_{(x3)}$ is the area location impact index. And the a b c is the different weight, comes from the expert analysis, $a = 0.5, b = 0.3, c = 0.2$.

Table 2 The road connectivity impact degree

Road type/impact	Directly connect with the core area	Connect sub-center with the core area	Connect peripheral area with sub-center	Others
Impact evaluation	10	8	6	3

Table 3 Evaluation of impact degree of road administration grade

Road grade/impact evaluation	Express way	Major road	Secondary roads	Slip road
Impact evaluation	10	8	6	3

4.1 Road Connectivity Index

Besides the location of the road section can have different impacts on the fast connection function of the city sub-center, the road connection function has a more direct impact on the fast connection function of the city sub-center. For example, for the two roads with the same administrative level within the city sub-center, the road directly reaching the administrative office core area has a more obvious impact on the fast passage of the administrative office core area than the road that cannot be directly reached. Therefore, the road connectivity function is selected as an index for the impact analysis of road congestion (Table 2).

4.2 Road Grade Impact Index

Highways and urban roads of different administrative levels have different traffic capacities, transportation functions and political significance. Under the same congestion degree, the impact of urban expressways is greater than that of branch roads, and the impact of expressways is greater than that of county and township roads. Therefore, the road administrative grade is selected as the index of road congestion impact analysis (Table 3).

4.3 Road Location Impact Index

According to the location and layout of the administrative office core area, the city sub-center and the peripheral control area, the roads in Tongzhou district can be divided into the internal roads of the administrative office core area, the internal roads of the city sub-center and the roads of the peripheral control area at present. roads in different locations have different main travel functions, for example, the internal roads of the administrative office core area mainly serve the short-distance business travel of public servants; Some roads in the administrative sub-center are

Table 4 Road location impact

Road location	Core are	Sub-center	Peripheral area
Impact evaluation	10	8	6

the “only way” for the peripheral control area to enter the administrative office core area. Some trunk roads and expressways in the peripheral control area are the main passages for other areas to enter the city’s sub-center. Therefore, the location of road sections is selected as an index for analyzing the impact of road congestion (Table 4).

5 Calculation of Road Congestion Impact Index in Urban Deputy Center

According to the above analysis model, calculate the impact index of road congestion in the urban sub-center and peripheral control areas. The calculation results are shown in Table 5.

Table 5 Ranking of road congestion impact index

Ranking	Road name	Congestion index
1	Xinhua west street	87.5
2	Xinhua south road	83.4
3	Xinhua west street auxiliary lane	79.8
4	Chaoyang north road	77.6
5	Guangqu road	77.1
6	Yongshun south street	68.9
7	Wuzhi xueyuan Road	66.4
8	Tonghui south road	64.3
9	Chezhan road	62.2
10	Chezhan road (near Xiying street)	58.4
11	Luyuan south street	55.6
12	Xiyuanting Street	55.4
13	Xiguan street	53.6
14	Furong east road auxiliary lane	53.4
15	Furong east road	52.9
16	Yile middle road	52.7
17	Longwang west road	51.4
18	Yaocun road	50.4
19	Xincang road	50
20	013Xi road	49.5

Comparing Tables 1 and 4, it can be seen that by calculating the traffic impact degree of the congested road sections, the ranking of road congestion has changed significantly, especially that many of the top 20 congested roads in Table 1 do not appear in Table 4. Therefore, according to the congested road section in Table 4, we can carry out a detailed field investigation on the road of Tongzhou sub center, and give the congestion analysis scheme.

6 Conclusion

Based on the analysis of traffic congestion in existing corridors, this project will allocate the traffic volume to be increased in the future according to factors such as relocation traffic volume and traffic proportion of each corridor, and finally determine the traffic volume of each corridor. Combined with the existing congestion situation, the congested road sections and congested nodes are given. It provides a reliable basis for future reconstruction.

Research on the Plan of Container Train Operation in Sea-Rail Intermodal Transportation Under the Uncertain Demand



Zhiyi Lei, Zhikuan Ma, Jéssica Saturnino Sabino de Sousa, and Yifan Zhu

Abstract The sea-rail intermodal container train is an important part of the sea-rail intermodal system. Its operation plan directly affects the detention time of sea rail intermodal containers in the port. In view of the characteristics of the large fluctuation of the container flow to the port, the Gauss hidden Markov model is employed to predict the container flow and a multi-objective mixed integer-programming model is established to minimize the detention cost of container and the cost of the trains. As an example, different scenario needs of Yangshan Port in Shanghai are forecasted through Gauss hidden Markov model, and the container train operation scheme under different scenarios is optimized for the terminal front loading and unloading container scheduling scheme.

Keywords Container terminal · Sea-rail intermodal transportation · Container train

1 Introduction

Economic globalization makes the trade between countries more and more frequent, and makes the demand of freightage between countries surge. Most multinational companies choose sea transportation to meet transportation needs. There are many port groups along the east coast of China. Goods from Europe, America and Southeast Asia tend to gather at the ports along the coast of China, and then evacuate to the inland of China by railway and highway. The convenient multimodal transport mode of sea rail combined transport has been paid more and more attention by the state.

The existing references on sea-rail intermodal transport system focuses more on the collaborative optimization of various subsystems and the prediction of coarse grain size of sea rail intermodal container. Corman et al. [1] defined the transportation between sea-rail intermodal transport systems as container transportation between terminals and container railway transportation, and established a mathematical model for port container transportation planning. An integer linear programming model was

Z. Lei (✉) · Z. Ma · J. S. Sabino de Sousa · Y. Zhu
Hohai University, Nanjing 210098, China
e-mail: lei_zhiyi@163.com

established to plan the cross transportation of container terminal, container terminal yard, port railway station and railway to hinterland. A tabu search algorithm was proposed, and its practicability was verified in practical cases and different demand scenarios. Schepler et al. [2] proposed an optimization model for multi terminal and multi railway container ports. Starting from the scheduling of ships, trains and trucks on the terminal, the transportation process of containers between terminals is reasonably optimized, and the weighted turnover time of containers in ports is minimized. Tierney et al. [3] proposed an integer programming model to express the flow process of vehicles and containers in sea rail intermodal transport system by using time-space diagram, which can provide important information for port scheduling and decision-making by reducing the transportation time of vehicles and containers in the system.

Wu and Liu [4] used the grey theory and radial basis function neural network to predict the development trend of sea-rail intermodal container of Ningbo port in the next two years, taking the container volume in recent six years as the original data. Wei [5] predicted the container volume of sea-rail intermodal transport in Fujian Province based on the economic growth rate, the throughput of each port, the regional GDP and the amount of foreign direct investment. Lu [6] used two prediction models of Grey Markov chain model and grey BP neural network model to predict the container sea rail combined transport volume of Dalian port in the next four years.

Aiming at the problem of train operation of sea-rail intermodal container, with the characteristics of container connection process in sea-rail intermodal transportation system, a multi-objective programming model is established to minimize the detention cost of containers in the port and the operating cost of trains. The difficulty of this problem lies in the fact that the quantity of sea-rail intermodal containers arriving at the port fluctuates greatly due to the Markov characteristics of container flow, which challenges how to adjust the number and frequency of train formation in the operation scheme of container trains. The detention time of container in port and the cost of train operation are optimized, and as well as the unloading sequence of sea-rail intermodal container and the number and frequency of train operation. Taking Yangshan Port of Shanghai as an example, the genetic algorithm is employed to solve the model.

2 Problem Description

In this paper, the problem of train operation in sea rail intermodal transportation is defined as the transshipment of imported containers by multiple ships and trains with different destinations in a research period. The research period is scheduled to be the length of the berthing plan designated by the port, which is generally 7 days; the ship is mainly responsible for the unloading of containers, and the trains are mainly responsible for the loading of containers. The research contents include the unloading sequence of containers requiring sea-rail intermodal, the operation scheme of trains, and the source of container flow on the train. The premise of the whole sea

rail intermodal operation process is to ensure that the containers of the berthing ships can be loaded on the trains with different destinations. In addition to the above rigid constraints, the whole sea-rail intermodal connection process is also easily affected by the arrival container flow and the train operation scheme.

3 Establishment of Basic Model

3.1 Minimizing the Detention Time of Sea Rail Intermodal Containers in Ports

The problem studied in this paper is defined as the operation of imported sea-rail intermodal trains. S means ships (the set of ships, the identify number of ships is expressed by s), N means trains (the set of trains, the identify number of trains is n), and D means the destination of those individual trains (the destination set of trains, the identify number of trains is expressed by d). The other parameters of the model are shown in Table 1. The paper assumes that the detention time of sea-rail

Table 1 Some parameters

h_i	The number of container for sea-rail intermodal transport of the i th ship (TEU)
T'_{ij}	The departure time of the j th train corresponding to the sea-rail intermodal container of the i th ship
a_{ij}	The beginning unloading time of the i th ship corresponding to the j th train
b_{ij}	The end unloading time of the i th ship corresponding to the j th train
q_j	The initial stored containers belong to the j th train in the yard
q'_j	The number of sea-rail intermodal containers belonging to the j th train that are loaded into the yard by trucks
q''_j	The number of sea-rail intermodal containers belonging to the j th train taken out of the yard by the container trucks
Q	Maximum storage capacity of storage yard
t_{ij}	Loading and unloading operation time
c_x	Detention cost coefficient of sea-rail intermodal container in port
c_a	Fixed cost coefficient of train operation
c_d	Variable cost coefficient of train operation
c_d	Under Axle cost coefficient of train operation
\bar{m}	The average mass of a standard container in tons
\bar{t}	Average storage time of container in yard
n_{if}	Maximum transshipment time of containers of the i th ship to f destination
m_{fa}	Maximum number of marshalling trains to f destination
m_{fb}	Minimum number of marshalling trains to f destination

intermodal container is from the time when the ship starts to unload to the time when the corresponding destination train starts to run, which mainly includes the operation time of shore crane, the storage time of wharf yard, the transportation time of container in the port and the assembly time of container.

For the port side, reducing the detention time of containers in the port can relieve the stacking pressure on the terminal yard and shorten the connection time of sea-rail intermodal transportation. To shorten the detention time of containers in the port reasonably, it is necessary to adjust the unloading sequence of containers on the ship on the basis of ensuring that the containers on each ship can match the corresponding destination trains. For example, there are sea-rail intermodal containers with destinations of d_1, d_2 and d_3 on the ship of s . The corresponding unloading time is T_1, T_2 and T_3 . The first container train is n_1, n_2 and n_3 corresponding to the destinations of d_1, d_2 and d_3 after the ship starts unloading. The corresponding departure time of each train is t_1, t_2 and t_3 . If $t_3 < t_1 < t_2$, the unloading time of containers with different destinations at this time shall meet the requirements $T_3 < T_1 < T_2$. In this way, the detention time of containers in the port can be reduced. The above plan is mainly to reduce the detention time of sea-rail intermodal containers in the port, by detaining the containers that do not need to be transferred by train on the ship before the ship leaves the port, and sending the containers that need to be transshipped to the port station for marshalling as soon as possible.

Therefore, from the point of view of the port, it is necessary to optimize the start of sea-rail intermodal container and reduce the detention cost of container in the port from the perspective of ship unloading time. In this paper, the detention cost of sea-rail intermodal container in port is taken as an important parameter to measure the connection reliability of sea-rail intermodal transportation system, and the objective function is to minimize the detention cost of container. The objective function is:

$$\min \omega_1 = c_x \cdot \sum_{i=1}^s \sum_{j=1}^n (T'_{ij} - T_{ij}) \cdot x_{ij} \tag{1}$$

The decision variables of the model are expressed as follows (Table 2).

Table 2 The decision variables

T_{ij}	The start of unloading time of the sea-rail intermodal containers of the i th ship corresponding to the j th train
x_{ij}	Container quantity of the j th train from the i th ship
m_{fj}	The number of marshalling of the j th train to f destination
x_{fj}	$x_{fj} = \begin{cases} 0 & \text{The } j\text{th train to } f \text{ destination will not operate} \\ 1 & \text{The } j\text{th train to } f \text{ destination will operate} \end{cases}$

3.2 Minimizing the Operating Cost of Trains

For the railway side, the operating cost of the train is an important factor in the design of the train operation scheme. When the number of sea-rail intermodal containers needed to be transferred by the train is too small, the unchanged operation schedule will cause waste of locomotive transport capacity, increase the under axle ratio of running trains.

For example, in a certain period, there are so few sea-rail intermodal containers that need to be transferred to d destinations in a certain day. According to the original plan, the containers can be loaded into the train of d -destination issue on that day, and the minimum number of marshalling of the train is 25. However, in the actual situation, the scheduled trains on the same day can choose not to run, and put the containers to be transferred in the yard, and the next train to d destinations in the study period will be transferred because the number of containers to be transferred is too small. For the port side, the above scheme increases the detention cost of containers in the port, but for the railway side, it may reduce the operating cost of trains and reduce the under axle ratio of trains.

Therefore, the economic benefits of railway departments should be considered in the design of train operation in the sea-rail intermodal system. In the design of train operation scheme, it is necessary to optimize the train operation frequency and the number of train groups reasonably, which can effectively reduce the cost of train operation and the under axle ratio of train. The corresponding objective function is as follows:

$$\begin{aligned} \min \omega_2 = & \sum_{f \in F} \sum_{j \in N} x_{fj} \cdot c_a + \sum_{f \in F} \sum_{j \in N} x_{fj} \cdot c_b \cdot m_{fj} \\ & + \sum_{f \in F} \sum_{j \in N} (m_{fa} - m_{fj}) \cdot c_d \\ & + (k_1 + k_2) \cdot \bar{m} \cdot \bar{t} \cdot \left(q_j + \sum_{j \in N} q'_j - \sum_{j \in N} q''_j \right) \end{aligned} \tag{2}$$

k_1 represents the storage cost of containers in the yard.

Refers to the loss of value of goods per hour, after the container is detained.

The constraints of the two objective functions are as follows:

1. The unloading time of sea-rail intermodal container must be within the loading and unloading time of the ship:

$$a_{ij} < T_{ij} < b_{ij} \quad \forall i \in S, \forall j \in N \tag{3}$$

2. The total unloading time of sea-rail intermodal container must be within the loading and unloading time range of the ship

$$a_{ij} < T_{ij} + x_{ij} \cdot t_{ij} < b_{ij} \quad \forall i \in S, \forall j \in N \quad (4)$$

3. The number of sea-rail intermodal containers on the j th train of f destination corresponding to the i th ship shall not exceed the number of sea-rail intermodal containers transported to f direction.

$$0 \leq x_{ij} \leq x_{if} \quad \forall i \in M, \forall j \in f \quad (5)$$

4. The amount of sea-rail intermodal containers loaded by a container train shall be equal to the maximum container capacity of the train:

$$\sum_{i \in S} x_{ij} = m_{fj} \cdot x_{fj} \quad \forall f \in F, \forall i \in S, j \in N \quad (6)$$

5. The number of sea-rail intermodal containers from trains on the i th ship shall be equal to all the sea-rail intermodal containers of the i th ship:

$$\sum_{j \in N} x_{ij} = h_i \quad \forall i \in S, \forall j \in N \quad (7)$$

6. The number of sea-rail intermodal containers loaded on the i th ship from all trains to f direction shall not exceed the number of sea-rail intermodal containers transported to the f direction:

$$\sum_{j \in f} x_{ij} \leq x_{if} \quad \forall i \in S, \forall j \in N \quad (8)$$

7. For sea-rail intermodal containers to f direction, the transshipment time of all containers must be within a certain transshipment time:

$$\sum_{j \in f} x_{ij} \cdot (T'_{ij} - T_{ij}) \leq n_{if} \quad (9)$$

8. The capacity constraints of storage yard is as follow:

$$q_j + \sum_{j \in N} q'_j - \sum_{j \in N} q''_j \leq Q \quad (10)$$

9. The time after unloading the container from the i th ship to the j th train shall not exceed the time when the train starts to issue:

$$T_{ij} + t_{ij} \times x_{ij} < T'_{ij} \quad (11)$$

10. The value constraints of variables are:

$$m_{fa} \leq m_{fj} \leq m_{fa} \quad m_{fj} \in N^+, \forall f \in F, \forall j \in N \quad (12)$$

$$x_{ij} \in N^+ \quad i \in S, j \in N \quad (13)$$

$$x_{fj} \in \{0, 1\} \quad (14)$$

The objective function (1) is to minimize the detention cost of sea rail intermodal container in the port, for the purpose to improve the connection reliability of various subsystems of sea-rail intermodal transportation, and to solve the unloading sequence of sea-rail intermodal containers on the ship; the objective function (2) is to minimize the operating cost of sea-rail intermodal trains, with the purpose of the economic interests of the railway sector, which consider of the fixed cost and variable cost. Reducing the detention time of sea-rail intermodal container in the port means that the operation frequency of the train and the number of train groups will be increased, which will lead to the increase of the objective function (2). Therefore, there are contradictions in the optimization of the two objective functions, and the two objective functions represent the interests of the port side and railway side respectively.

4 Establishment of Optimization Model

The arrival container flow is obviously one of the important factors that affect the reliability of sea-rail intermodal transportation. Too large container flow leads to cargo pressure, and too small container flow leads to surplus capacity of trains. According to the relevant literature [7], the fluctuation of arrival container flow is essentially a Markov characteristic, which will impact on the transportation efficiency of the system and destroy the reliability of the system connection.

In view of the Markov characteristics of the container flow, the change process of the arrival container flow can be approximately equal to the Markov process. In this paper, the probability distribution of the box flow in the next period is determined by the Gaussian hidden Markov model [8] (GM-HMM), and the uncertainty problem is transformed into the certainty problem under the discrete probability. To a certain extent, it can also reflect the changes and influences of the future container flow, and improves the robustness of the model, so that the whole system can also obtain a reasonable scheme under different sea rail intermodal container transport demand.

4.1 Hidden Markov Model for Sea-Rail Intermodal Transportation System

The parameters of hidden Markov model are explained here. N means the total number of states in the model. The number of sea-rail intermodal containers o in the period analyzed in this paper is a continuous variable, and there are 3 hidden states, respectively, the container quantity is in the rising state, falling state and stable state.

1. o_t means observed value. The observed values are the daily container flow, the growth rate of container flow and the growth rate of container flow.
2. π means probability vector of initial state. Considering the change law of container flow, the default initial state is stable state.
3. A means state transition matrix. It expresses the change process of containers. It is noted that the change state of container at the time t is state q_i , and next time ($T + 1$) is q_j , then the state transition probability is defined as $a_{ij} = P(i_{t+1} = q_j | i_t = q_i)$, and the state transition matrix is formed.
4. For Gaussian hidden Markov model, there is no traditional observation state matrix. Being one-dimensional, the observed value of sea-rail intermodal containers can be considered obeys a Gaussian distribution. For each state q_i , the sea-rail intermodal container is:

$$b_n(O) \sim N(\mu_n, \sigma_n) \quad n = 1, 2, \dots, N \quad (15)$$

And the model parameters needed by Gaussian hidden Markov are as follows:

$$\lambda = (A, \pi, \mu, \sigma) \quad (16)$$

4.2 Parameter Learning and Model Prediction

In the process of container flow fluctuation, the general situation of internal system can not be observed but the size of container flow. The corresponding data of container flow is the training data of the model. The GM-HMM of sea-rail intermodal container is established, and the parameter $\lambda = (A, \pi, \mu, \sigma)$ of the model is trained by the algorithm (Baum Welch algorithm [9]). In the prediction stage, the parameters of GM-HMM learned in the previous link are known. According to the Viterbi algorithm, the state sequence of GM-HMM can be obtained. According to the state transition matrix A , the probability of each state in the period of $T + 1$ can be calculated, that is $P = \{P(i_{T+1} = q_1), P(i_{T+1} = q_2), P(i_{T+1} = q_3)\}$. Furthermore, it is known that the probability density function of periodic sea-rail intermodal container is as follows:

$$Z = \sum_{k=1}^3 P(i_{T+1} = q_k) \cdot N(\mu_k, \sigma_k) \tag{17}$$

4.3 Modeling Under Uncertain Demand

The Gaussian probability distribution function corresponding to the predicted state of the next cycle is defined as the probability distribution corresponding to the future scenario demand. Demand of each scenario corresponds to the possible transportation demand of sea-rail intermodal container in the next cycle. Based on the basic train operation model of sea rail intermodal transportation, the transportation demand under different scenarios is considered. If the container flow increases sharply in a certain scenario, it is considered that container trains need to increase the number of marshalling, thus the fixed cost of train operation will increase. If the container flow is reduced in a certain scenario, it is considered that the container trains need to leave under axle, and the corresponding under Axle cost will be generated.

So the model should be optimized under uncertain demand with the following parameters:

R : Demand set of sea-rail intermodal container predicted by Gaussian hidden Markov model, $\forall r \in R$.

μ_r, σ_r^2 : Mean and variance of Gaussian distribution under scenario r .

p_r : Probability of occurrence predicted by Gaussian hidden Markov model under scenario r .

$m_{f,r,j}$: Marshalling number of the j th container trains to f destination under scenario r .

1. Objective Function

The objective function consists of two parts. The first part has the same operating cost as the model under deterministic demand. Considering the total cost of sea-rail intermodal transportation, the function is as follows:

$$\begin{aligned} \min \omega_2 = & \sum_{f \in F} \sum_{j \in N} x_{fj} \cdot c_a + \sum_{f \in F} \sum_{j \in N} x_{fj} \cdot c_b \cdot m_{fj} \\ & + \sum_{f \in F} \sum_{j \in N} (m_{fa} - m_{fj}) \cdot c_d \\ & + (k_1 + k_2) \cdot \bar{m} \cdot \bar{t} \cdot (q_j + \sum_{j \in N} q'_j - \sum_{j \in N} q''_j) \end{aligned} \tag{18}$$

In the second part, different situations in the future and the fluctuation of sea-rail intermodal containers are considered. The number of marshalling trains will affect the variable cost of trains. the change of cost therefore should be considered as follows:

$$\min \omega_4 = \sum_{r \in R} \sum_{f \in F} \sum_{j \in N} |m_{fjr} - m_{fj}| \cdot c_b \quad (19)$$

The total objective function is as follows:

$$\begin{aligned} \min \omega_2 = & \sum_{f \in F} \sum_{j \in N} x_{fj} \cdot c_a + \sum_{f \in F} \sum_{j \in N} x_{fj} \cdot c_b \cdot m_{fj} \\ & + \sum_{f \in F} \sum_{j \in N} (m_{fa} - m_{fj}) \cdot c_d \\ & + (k_1 + k_2) \cdot \bar{m} \cdot \bar{t} \cdot (q_j + \sum_{j \in N} q'_j - \sum_{j \in N} q''_j) \\ & + \sum_{r \in R} \sum_{f \in F} \sum_{j \in N} |m_{fjr} - m_{fj}| \cdot c_b \end{aligned} \quad (20)$$

2. Constraint conditions

The constraints mainly consider the meet of transportation demand under different scenarios.

$$\sum_{f \in F} \sum_{j \in N} m_{fjr} \geq \mu_r \quad (21)$$

The decision variable is a positive integer.

$$m_{fjr} \in N^+ \quad \forall f \in F, r \in R, j \in N \quad (22)$$

4.4 Algorithm Design

Bi-level programming model is NP hard problem. When there are too many ships and trains in the period, it is difficult to get the optimal solution. In general, intelligent algorithm is employed to solve the problem.

1. Program:

The genetic algorithm in this paper is binary coding. In the above model, the number and frequency of trains are determined, and then the unloading time of containers on the ship and the container flow on the train are generated, and the time and container flow are transferred into binary.

2. Algorithm design process

Step 1: generate N initial solutions randomly, and generate initial population in this step;

- Step 2: set the maximum evolution algebra, crossover and mutation probability;
- Step 3: turn to the lower objective function to solve the operating cost of the train;
- Step 4: the upper objective function is solved according to the marshalling number of trains and operation frequency;
- Step 5: calculate the fitness of each chromosome and select it;
- Step 6: carry out cross mutation operation according to the crossover and mutation probability set in step 2. Judge whether the conditions for ending evolution are met. If the conditions for ending evolution are not met, turn to step 3; otherwise, go to step 7 directly;
- Step 7: output the results.

5 Case Analysis

Taking the Yangshan Port of Shanghai as an example. There are 9 berths, 15 wharf, 24,386 TEU of plane space of yard, 45 gantry cranes and 220 trucks. The results of the model are shown in Table 3.

Through calculation, we can get the optimal number and frequency of trains corresponding to berthing plan. If the number of train groups is 0, it means that the train will not run. Compared with the traditional mode, the scheme is more flexible and can reduce the detention time of containers in the port. The average detention time of containers in the above cases is 8.3 h, which realizes “seamless connection” to a certain extent. Compared with the traditional fixed-point assembly operation scheme, the advantages of this scheme are in the case of less arriving ships and less demand for container transportation, the extra cost caused by some trains can be avoided.

It’s shown from the above table that the container flow fluctuates. The trains with the first cycle tend to be under Axle travel, and the fluctuation of container flow has little impact on it. The number of marshalling should be planned according to the minimum number of marshalling, and the no-fixed trains will not be operated. The fluctuation of container flow has a great impact on the later trains because the fluctuation of container flow brings a large number of containers in the yard, which increases the pressure on the operation of the later cycle trains. Therefore, the marshalling number of trains at the following cycle should be planned according to the above table.

6 Conclusion

In view of the complex operation process of sea-rail intermodal container, a multi-objective integer programming model with the shortest detention time of containers in the port and the minimum operating cost is established, from the view of port side and railway side. Considering the Markov characteristics of container flow, the

Table 3 Optimal marshalling quantity for each scenario

Train	Running time	Number of marshalling
Luchao port-North Zhuzhou	2019/12/25 17:33	25
Luchao port-Jishan	2019/12/25 20:00	25
Luchao port-Dahongmen	2019/12/26 0:35	0
Luchao port-Turpan	2019/12/26 2:40	0
Luchao port-West Wangjiaying	2019/12/26 5:59	25
Luchao port-Turpan	2019/12/27 2:40	36
Luchao port-North Zhuzhou	2019/12/27 17:33	25
Luchao port-Jishan	2019/12/27 20:00	33
Luchao port-Dahongmen	2019/12/28 0:35	25
Luchao port-Turpan	2019/12/28 2:40	0
Luchao port-West Wangjiaying	2019/12/28 5:59	25
Luchao port-Turpan	2019/12/29 2:40	50
Luchao port-North Zhuzhou	2019/12/29 17:33	29
Luchao port-Jishan	2019/12/29 20:00	43
Luchao port-Dahongmen	2019/12/30 0:35	32
Luchao port-Turpan	2019/12/30 2:40	43
Luchao port-West Wangjiaying	2019/12/30 5:59	44

container transportation demand is forecasted in different situations by Gaussian hidden Markov model, and the basic model is optimized. The genetic algorithm is used to solve the model, which provides decision support for the loading and unloading operation of the wharf front and the train operation scheme design of the port station.

Compared with the existing research, the contribution of this paper lies in: considering the influence of container flow on the operation of container trains, an operation resource allocation model is established, which takes the connection reliability of sea-rail intermodal transportation system and the freight turnover efficiency as the goal. This study provides a new idea to solve the problem of how to reasonably design the operation scheme when the container flow fluctuates.

References

1. Hu Q, Corman F, Wiegmanns B et al (2018) A tabu search algorithm to solve the integrated planning of container on an inter-terminal network connected with a hinterland rail network. *Transp Res Part C: Emerg Technol* 91:15–36
2. Schepler X, Balev S, Michel S et al (2017) Global planning in a multi-terminal and multi-modal maritime container port. *Transp Res Part E: Logist Transp Rev* 100:38–62
3. Tierney K, Voß S, Stahlbock R (2014) A mathematical model of inter-terminal transportation. *Eur J Oper Res* 235(2):448–460
4. Wu H, Liu G (2015) Container sea-rail transport volume forecasting of Ningbo port based on combination forecasting model. In: *International conference on advances in energy, environment and chemical engineering*. Atlantis Press
5. Wei H (2014) Study on Forecasting of Sea-rail Inter-modal Transportation: in the case of ports in Fujian. *Logist Technol* (11):308–311
6. Lu Y (2018) Forecast of sea-rail intermodal container volume of Dalian Port based on combination forecasting model. *China Market* (27):21–24
7. Cheng J (2016) Research on connection reliability of sea-rail intermodal transport based on dynamic container train operation plan. Beijing Jiaotong University
8. Ji M, Yan Y, Zhu H (2018) Railway pricing and operation scheduling in rail-sea intermodal transportation. *Chin J Manage Sci* 26(7):159–169
9. Gauvain, J-L, Lee C-H et al (1994) Maximum a posteriori estimation for multivariate Gaussian mixture observations of Markov chains. *IEEE Trans Speech Audio Process*

Research on Setting of Warning Signs for Highway Tunnels in High Altitude Areas



Jian-you Zhao, Wei-guo Xie, Wan-li Han, and Wan-jiang Guo

Abstract In order to study the driving conditions of drivers on highway and tunnel sections in high altitude areas, took some sections of Gong Yu highway as the research object, and the driving speeds of the subjects were collected on sections at different altitudes to verify the influence of road environment on drivers' behavior characteristics in high altitude area. Aimed at the problem of irregular installation of traffic facilities on highway and tunnel sections in high altitude areas, and considered the distribution and species of animals along highway in high altitude areas, measures have been taken to install animal crossing warning signs. Based on the visual recognition characteristics of drivers, construct a model of the warning signs pre-distance for highway tunnels in high altitude areas, and get the speed limit signs pre-distance reference table.

Keyword Highway · Tunnel warning signs · Driver characteristics · Plateau high altitude · Pre-distance · U491 · A

CLC Number U491 · A

Document Code A

High altitude areas such as Tibet Autonomous Region, Qinghai Province and some Sichuan Province, their special natural geographical location, harsh environmental conditions, complex geological conditions and variable climate, have a greater impact on drivers (especially those the first time on the plateau highway). It often causes fatigue, slow response and speeding in the absence of oxygen and monotonous road environment [1, 2]. In the highland area, driving 1 to 3 h will appear obvious fatigue symptoms, leading to the occurrence of traffic accidents [3]. At present, domestic scientific research workers have done a lot of research on driving safe driving in the highland area. In view of the environmental characteristics of low-pressure low

J. Zhao · W. Xie (✉) · W. Han · W. Guo
School of Automobile, Chang'an University, Xi'an 710064, Shaanxi, China
e-mail: 15634211069@163.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_43

563

oxygen, strong ultraviolet radiation and high demeanor in the plateau, Eli analyzed the main factors affecting the safety of driving and established the structure of the driver's subjective feeling index [3]. Zhu et al. [4] statistically summarize the 986 more complete accident records on the main roads on the Qinghai-Tibet Plateau in China, and made a comprehensive analysis of the causes of the accidents and the distribution of space and space. In addition, the safety of tunnel traffic in the plateau area is also worthy of attention. Because the highway tunnel is a special road structure that is semi-isolated from the outside world, it has the characteristics of relatively small space, strong closure, large difference between the light environment inside and outside the tunnel, low brightness level in the tunnel and slow diffusion of soot, and so on, once the traffic accident occurs in the tunnel, it is easy to cause secondary accidents, resulting in difficulties in rescue treatment, economic loss and great social impact. One of the main reasons for the frequent tunnel accidents is that the difference in the light environment of the tunnel mouth causes the driver's visual recognition ability to decrease, and the necessary braking and deceleration measures can not be taken before entering the tunnel, which seriously affects the safety of the road. Large number of real vehicle experimental data analysis recorded by the eye movement device, Pan et al. [5] put forward that the main reason for the tunnel entrance section accident is that the driver's vision can not adapt to the sharp change of the brightness of the tunnel entrance section, study the change law of the driver's pupil in the whole process, and obtain the road tunnel entrance lighting safety evaluation method. Through the real vehicle test data of many tunnel entrances and exits, Du [6] constructs the correspondence between the illumination of the tunnel entrance and exit and the area of the driver's pupil, based on the driving person's pupil area and its changing speed, obtained the driving safety evaluation index, and put forward the range of the road tunnel entrance and exit lighting transitio. Based on the experimental data and analysis of the real vehicle of the mountain highway, Hu [7] constructs the regression model of the change rate, running speed and brightness reduction coefficient of the pupil area, and proposed that the threshold of safe and comfortable driving drivers was 20% of the pupil area change rate. Huang [8] constructs the construction of a model of the tunnel entrance sunshade shrink, and the practical application effect of the awning was verified. More research work on the impact of light on drivers at the tunnel mouth of the plains area, the driver's driving vision has been improved. But it can also cause some drivers to relax their vigilance when the light improves but does not completely clarify the road conditions ahead, causing traffic accidents to be caused by speed before entering the tunnel. At present, the study of highway tunnel front warning signs in high altitude areas is relatively brief, the relevant front warning signs are placed unreasonable, coupled with the effects of special environmental conditions in high altitude areas, some slow-down brake warning signs can not fully attract the attention of drivers, resulting in accidents. In this paper, with a more representative high altitude regional highway, Qinghai Province, Hainan Tibetan Autonomous Prefecture, the Republic County to Yushu Tibetan Autonomous Prefecture, the Jigu Town Highway (the Gong Yu highway) as the research object, the driver driving safety, the surrounding driving environment and the tunnel section of the warning signs set the problem of analysis,

found that in the high altitude area driving driver's consciousness will be affected, the road environment is easy to cause driver fatigue, there are wild animals, causing greater disturbance to drivers, And before entering the tunnel, there are no significant pre-warning signs, which have great hidden dangers to the safety of the road. In view of these problems, it is proposed to set up significant animal walking warning signs and passageways in the ordinary wild animal sections, and based on the characteristics of the driver's recognition, to establish a high altitude highway tunnel speed limit warning signs pre-distance model, in order to provide a certain reference and ideas for the safe driving of highways in high altitude areas.

1 Research Plan

The pilot with rich driving experience was selected, and corrective vision was above 5.0. Taking into account safety factors, the trial finally selected 3 healthy drivers of different ages and ages, all male, aged 28–45 years old, driving age of 4–10 years. One of the drivers had experience in highland driving, two drivers were first to the plateau, and all drivers had driven more than 2000 km in the previous year.

The test takes the real vehicle test, uses 2 high-definition dashcams to collect the vehicle's operating status and driver behavior characteristics, and uses the handheld Garmin GPS to collect the driving speed, altitude changes and other data.

2 Investigative Forensics and Data Analysis

2.1 *The Influence of the Road Environment in High Altitude Area on the Characteristics of the Driver's Behavior*

2.1.1 Effects on the Driver's Behavior

The driving speed of the subjects was collected at different altitude sections of the GongYu Highway to verify the influence of the road environment in the high altitude area on the characteristics of the driver's behavior. Using handheld GPS speed positioning, the driver's speed and driving time are collected on three different altitude sections: 3000–3500 m, 3500–4000 m and 4000–4500 m. At the same time, during the test process, the driver's estimate of the current driving speed is recorded through inquiry. Examine the characteristics of the driver's speed estimation of moving objects and the anxiety of individual personality. A total of 65 segment data were collected, excluding 16 segments with large errors. As shown in Table 1, the subjects are the vehicle's operating speed and vehicle speed estimates at three different altitude sections.

Table 1 Vehicle speed and estimated vehicle speed of subjects at three different altitudes

Participants	Driver A/(km·h ⁻¹)			Driver B/(km·h ⁻¹)			Driver C/(km·h ⁻¹)		
	Average speed	Estimated speed	Estimation error	Average speed	Estimated speed	Estimation error	Average speed	Estimated speed	Estimation error
3000–3500 m	79.6	60	9.6	98	80	18	82.5	70	12.5
3500–4000 m	105	70	35	91	70	21	72.7	60	12.7
4000–4500 m	108	75	33	89.9	70	19.9	92.8	70	22.8

Table 2 Drivers' response delays at different altitudes

Altitudes	2000 m	3000 m	4000 m
Response delays/s	0. 77	0. 85	0. 89
Standard deviation	0.1010	0.1489	0.1123

It can be found that with the elevation of the rise, the average driving speed of the driver A and C have increased trend, and the speed estimation error gradually increased, its average driving speed is not between 72.7 and 108 km·h⁻¹, the speed estimation error value is 9.6–35 km·h⁻¹, some sections are higher than the design speed of 80 km·h⁻¹, more than the design speed of 3.1% to 35%. The average driving speed of driver B with the increase of altitude has a decreasing trend, which may be related to driver B often driving at high altitude, but its overall speed is high, the average speed is between 89.9 and 98 km·h⁻¹, are higher than the design speed of 80 km·h⁻¹, more than the design speed 12.4–22.5%, speed estimate error 18–21 km·h⁻¹, At an altitude of 4000–4500 m, the average driving speed of the three reached 90 km·h⁻¹, the highest speed 119 km·h⁻¹, the driver felt fatigue, slow response and breathing difficulties, and have mild dizziness symptoms, speed estimate error value is too large brings about speeding. This is basically consistent with the results of Wang study [9] on the speed of the Qinghai-Tibet Highway with an altitude of 4 km as the critical point, and the basic expected speed of 120 km·h⁻¹ and Feng Bingbing's study [10] of the air oxygen content of the road speed distribution law in Qinghai Province which the results of the study of the speed of the vehicle change from 60 to 90 km·h⁻¹. It is comprehensively verified that highway drivers at high altitudes are more affected by the environment than when driving on plain highways.

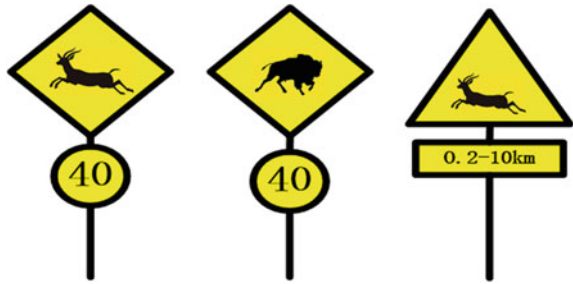
2.1.2 Effects on the Driver's Reaction Time

The experiment proves that when the driver is in the highland environment, the environmental influence of the plateau is relatively large. In the field test study of the physiological and psychological characteristics of the highland highway driver, Eli [11] collected the physiological and psychological effective sample data of 1079 drivers, and the study showed that with the increase of altitude, the reaction time increased compared to normal, increasing the time for drivers to cope with potential hazards and making them prone to collision accidents. As shown in Table 2, drivers' response delays at different altitudes.

2.1.3 Lateral Road Interference

High altitude highways are dotted with pastoral areas and wildlife nature reserves along the way, and drivers are prone to fatigue, slow response, low accuracy of speed estimation of moving objects, speeding, and more traffic accidents to promote

Fig. 1 Animal passing warning signs on highways in high altitude areas



collisions between vehicles and animals. On the eastern edge of the Qinghai-Tibet Plateau at an altitude of 3400 m, up to 5800 wild animals were crushed to death. An average of 1.9 animals, or 19 per m, were surveyed on 110 sample lines. It is also found that most of the high-altitude highway animal passages used cement and most of the animal passages and surrounding environment of the difference is large, low utilization rate, and lack of corresponding animal travel warning signs. The main distribution of animals in the pastoral areas through the Gong Yu highway are yaks, sheep, horses and sheepdogs, and there are often animals inside the highway guardrails, and there are no warning signs for animals to travel along the highway.

Animal passing warning signs can be set up to alert drivers to deceleration before vehicle-animal collisions occur, pastoral roads, access to pastoral areas and nature reserves. Different animals through the warning sign pattern can play a different warning effect, the study shows that the animals on the warning signs show the trend of movement when the impact on the driver is greater, the effect is better [12]. Taking into account the above factors, combined with the distribution and type of animals along the highways in high altitude areas, animal passing warning signs on highways in high altitude areas are shown in Fig. 1.

2.2 The Black-and-White Hole Effect of Tunnels in High Altitudes

Most high altitude highway tunnels are operating in low-light conditions. During the day, the tunnel entrance section is too sharp, resulting in the tunnel entrance driver dark adaptation difficulties, easy to induce the “black hole effect”, resulting in visual impairment, it is difficult to distinguish the traffic information inside the tunnel. At present, the main measures taken to improve the “black hole effect” of road tunnels in various provinces are to strengthen the lighting of the tunnel entrance section. Other measures include: the hole is designed as a “bamboo cut”, reflectors on both sides of the tunnel wall, a light-reducing sun visor at the hole, and a light-emitting paint. However, the investment and daily operating costs of these methods are high, and it is difficult to implement for some areas with low traffic flow. In the tunnel exit section, the phenomenon of glare is obvious, the driver is very prone to

Fig. 2 Black hole effect



Fig. 3 White hole effect



the feeling of vertigo, that is, the tunnel exit “white hole” effect, prone to tunnel exit vehicle tailing accidents, but according to the national norms set lighting energy consumption is high, and the Central and Western highway tunnel generally small traffic volume, the total operating income is far less than the tunnel operating lighting costs, operating units can not afford huge tunnel lighting costs, adopt long-term low-illuminated strategy (Fig. 2 and 3).

2.3 The Setting of Traffic Warning Signs for the Tunnel Section

The biggest problem with the setting of highway traffic signs at high altitudes is the lack of necessary pre-warning and post-confirmation signs. The setting of the pre-notice and post-confirmation signs enables the driver to prepare in advance at a certain distance in front of the pastoral area, tunnel crossing and intersection. Especially in the oxygen-deprived driving environment of the plateau, the driver’s attention, sensitivity and accuracy are significantly reduced, and the reaction time is extended, so that there is not enough time to react and operate. Drivers not notice the traffic signs, resulting in the loss of traffic information, the current high altitude highways generally do not set up the relevant warning signs.

As the current standardise is mainly for highway spoil, for the terrain is complex, the environment is relatively harsh special areas, especially for the highland mountain highway, the installation of traffic facilities does not have standard norms, there are some defects in the configuration of facilities. Taking into account the local topography, topography, elevation, road line, landscape environment along the route, humanities and other factors that may affect vehicles and drivers, taking into account the existing national industry standards. Optimize the configuration of security facilities, especially in tunnel sections. Based on the investigation and experiment, carry out the modeling of the front distance of the highway tunnel warning signs in high altitude areas.

3 The Pre-distance Model for Highway Tunnels Warning Signs in High Altitude Areas

The purpose of traffic warning signs is to remind vehicles or pedestrians to pay attention to dangerous places. There is sufficient time to take necessary actions when the danger is reached to ensure safe driving [13]. The process of driver’s recognition of traffic warning signs can be divided into 4 stages of “discovery, recognition, judgment, action” [14]. As shown in Fig. 4, the driver’s visualization of the warning sign. When the vehicle reaches the point A, it can identify the content of traffic warning signs. The distance from A to B is the reading distance of the driver S_1 . At this time, the driver takes action to judge (whether to brake). Then the driver receives the emergency stop signal, and his brain makes a braking instruction until the point

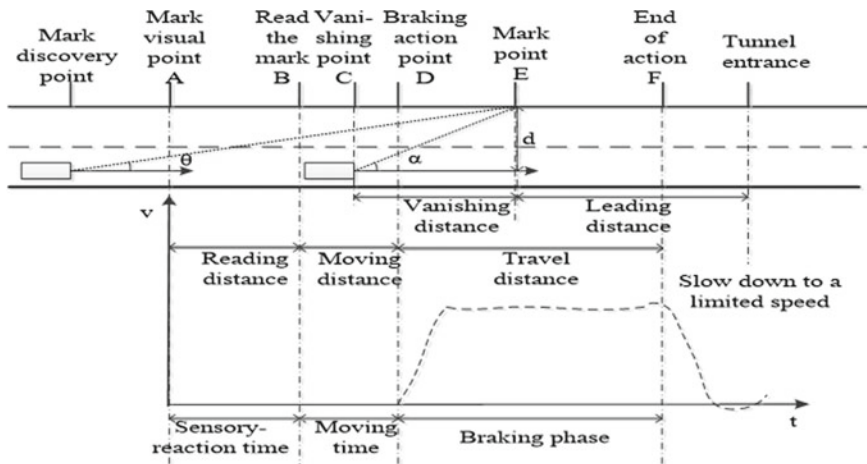


Fig. 4 Driver’s visual sign recognition process and corresponding braking deceleration and time curve

D driver moves his foot to the brake pedal to start action. This distance is the travel distance of the driver S_2 .

According to the driver’s dynamic visual characteristics, the driver’s perceptual search field of view has a certain range. When the angle α between the driver’s line of sight and the roadside sign at the sign setting point exceeds the viewing angle threshold θ , the sign will disappear from the field of view. The distance from the sign vanishing point C to the setting point E of warning signs is called vanishing distance m . The distance from the start of the action D to the end of the action F is the braking distance S_3 . The distance from the point A to the warning sign E is called the visual recognition distance S_0 . The distance from the reading point B of the sign to the sign E is the distance S_4 to the sign after reading the information. If the distance S_4 is shorter than the distance from the vanishing point C to the sign E, it means that the driver can’t accurately judge the content of the sign [15].

$$S_3 \leq D + S_4 - S_2 \tag{1}$$

$$S_3 \geq m = \frac{d}{\tan \theta} \tag{2}$$

where: θ is the driver’s viewing angle threshold, the general roadside sign is 15° , the overhead sign is 7° , d is the driver’s apparent height (1.2 m) to the side distance of the sign (roadside installation sign). According to the survey, the Gong Yu highway lane width is 3.75 m.

3.1 Calculation Model of the Visual Recognition Distance of Traffic Signs

There are regulations on the specifications of speed limit signs and recommended speed signs in the national standard. Analyze the relationship between the viewing distance S_0 , the actual height of the image h and driving speed. According to the principles of human factors engineering [16], The setting conditions of animal crossing warning signs and the size of pictures and texts are designed according to formula (3).

$$S_4 = f(h^*) \tag{3}$$

$$h^* = k_1 \times k_2 \times k_3$$

where: h^* is the height of the effective picture text; k_1 is the language type correction factor, take 1.2; k_2 is the text complexity correction factor: take 1 when the strokes of Chinese characters are less than 10, take 0.9 when 10–15 strokes; take 0.85 when

Table 3 The value of k_3

Speed/(km·h ⁻¹)	Walking	20	30	40	50	60	70	80	90	100
k_3	1	0.96	0.94	0.91	0.89	0.87	0.85	0.82	0.79	0.77

more than 15 strokes; k_3 is the driving speed correction factor, selected according to Table 3.

According to the experimental data of Japan Civil Engineering Research Institute [16]:

$$f(h^*) = 5.67h^* \tag{4}$$

The visual recognition distance calculation model is:

$$S_4 = 5.67h^* = 5.67 \times 1.2 \times 1 \times k_3 = 6.084k_3 \tag{5}$$

3.2 The Vehicle Deceleration Distance Calculation Model Based on the Driver’s Visual Characteristics.

According to the driver’s operating law, the driver reads the sign information, brakes the vehicle and eventually slows down in front of the speed control point. This process can be divided into two stages, the brake reaction phase and the brake phase [17].

- (1) Brake reaction phase. The brake reaction stage includes sensory-reaction time (corresponding to the visual recognition process to analyze the reading mark distance S_1) and moving time (corresponding to the visual recognition process to analyze the moving distance S_2). The reaction time of drivers in different altitude sections is 0.77–0.89 s, taking into account the extreme state, it takes 0.89 s. According to the principle of human factors engineering, the driver’s eye movement time is 0.09 s, the target selection time is 0.6 s, and the recognition time is 0.5 s. In the braking reaction phase, the driver’s feeling-reaction time is 2.08 s. Since the time of the driver’s moving phase (changing feet and depressing the brake pedal) is relatively small, generally 0.2–0.3 s, the limit state value is 0.3 s. At this stage, the vehicle is not braked and is running at a constant speed. Its running distance is shown in Eqs. (6) and (7):

$$S_1 = v_0 \times 2.08 \tag{6}$$

$$S_2 = v_0 \times 0.3 \tag{7}$$

- (2) Brake phase. This stage corresponds to the action distance of the visual recognition process and is divided into the brake transition stage and the smooth brake stage.

Brake transition stage. The deceleration of the vehicle changes from 0 to a certain deceleration a_0 , and the speed decreases from v_0 to v_1 . According to the principles of vehicle engineering [18], this deceleration value a_0 has a linear relationship with the initial speed of the vehicle, and its relationship model is as shown in Eq. (8):

$$a_0 = 0.524v_0 - 0.1215 \tag{8}$$

According to the empirical model of vehicle engineering, the vehicle braking transition phase time is generally 0.3 s, and it is considered that the deceleration obeys a linear change, then the speed at the end of the braking transition phase can be expressed by Eq. (9):

$$v_1 = v_0 - \int_0^{0.3} \frac{0.524v_0 - 0.1215}{0.3} t dt \tag{9}$$

The vehicle running distance in the brake transition phase is:

$$S_3 = v_0 \times 0.3 - \int_0^{0.3} \frac{1}{2} \left(\frac{0.524v_0 - 0.1215}{0.3} \right) t dt \tag{10}$$

Smooth brake stage. The vehicle travels at a constant deceleration from v_1 to the speed limit v_t specified by the speed control sign. The vehicle running distance in the steady braking stage is:

$$S_3 = \frac{v_1^2 - v_t^2}{2a_0} = \frac{(v_0 - \int_0^{0.3} (\frac{0.524v_0 - 0.1215}{0.3}) t dt)^2 - v_t^2}{2 \times (0.524v_0 - 0.1215)} \tag{11}$$

Based on the above two-stage distance calculation model, after the driver reads the tunnel warning sign, the vehicle deceleration distance model is:

$$S = S_1 + S_2 + S_3 = v_0 \times 2.68 - \int_0^{0.3} \frac{1}{2} \left(\frac{0.524v_0 - 0.1215}{0.3} \right) t dt + \frac{(v_0 - \int_0^{0.3} (\frac{0.524v_0 - 0.1215}{0.3}) t dt)^2 - v_t^2}{2 \times (0.524v_0 - 0.1215)} \tag{12}$$

3.3 Sign Pre-distance Calculation Model

From Eqs. (1) and (2), the conditions for setting the pre-distance of warning signs for animal crossing in high altitude areas are:

$$D \geq S - S_4 \tag{13}$$

$$d \leq S_3 \tan \theta \tag{14}$$

Substituting formula (5), formula (12), and formula (13):

$$D \geq v_0 \times 2.68 - \int_0^{0.3} \frac{1}{2} \left(\frac{0.524v_0 - 0.1215}{0.3} \right) t dt + \frac{(v_0 - \int_0^{0.3} \left(\frac{0.524v_0 - 0.1215}{0.3} \right) t dt)^2 - v_t^2}{2 \times (0.524v_0 - 0.1215)} - 6.084k_3 \tag{15}$$

$$d \leq \frac{(v_0 - \int_0^{0.3} \left(\frac{0.524v_0 - 0.1215}{0.3} \right) t dt)^2 - v_t^2}{2 \times (0.524v_0 - 0.1215)} \tan \theta \tag{16}$$

Bring in specific data for inspection, and you can get the pre-distance model of tunnel warning signs in high altitude areas to meet the conditions. Through the above calculation and analysis, in view of the speed limit of Gong Yu Highway, the leading distance is 767 m. The leading distance obtained according to the model is farther than the leading distance of general traffic speed limit signs in plain areas [17], This is due to the special geography and traffic environment in high-altitude areas, which have a greater impact on drivers [9, 11]. The calculation result also shows that the model fits the actual situation in high altitude areas. According to the above model, the distance between the speed limit sign and the recommended speed sign can be obtained by calculation when the initial speed is reduced to a certain speed limit value, as shown in Table 4.

Table 4 Speed limit signs pre-distance reference table

Driving speed /(km·h ⁻¹)	k ₃	Reduce to the following speed/(km·h ⁻¹) pre-distance/m					
		40	50	60	70	80	90
60	0.87	466.1	308.0	/	/	/	/
s70	0.85	621.3	486.7	331.6	/	/	/
80	0.82	767.0	649.8	514.6	355.0	/	/
90	0.79	906.3	802.5	682.8	541.3	378.0	/
100	0.77	1041.2	947.9	840.5	713.5	567.0	400.9

4 Conclusion

The vehicle speeds of the subjects were collected at different altitude sections of Gong Yu highway to verify the influence of the road environment in high altitude areas on the behavior characteristics of drivers. It is comprehensively verified that highway drivers in high altitude areas are more affected by the environment than when driving on plain highways. There are often wild animals on highways in high altitude areas that affect the safety of drivers. Animal crossing warning signs can be set up in places where vehicle and animal collisions frequently occur, pastoral roads, pastoral areas and nature reserves to remind drivers to slow down. At present, many current standards are mainly formulated for highways and tunnels. For special areas with complex terrain and relatively harsh environments, especially for mountain highways on plateaus, there are no standards and specifications for the installation of traffic facilities. There are some flaws. In the test drive experiment and data analysis, a model of the warning signs pre-distance for highway tunnels in high altitude areas is established, which aims to provide drivers with safe driving in high altitude environments and provide theoretical references for highway tunnel sections in high altitude areas. However, the current work is mainly concentrated on some sections of Gong Yu highway. Driver data collection is single, and only a simple analysis of the black and white hole effect of the tunnel is carried out. In the future, more extensive data collection and research will be carried out on highways and tunnel sections in high altitude areas.

Acknowledgments Supported by Foundation of Central University (310822171006).

References

1. Song C, Zhang S, Li J et al (1993). A study on the combined effect of hypoxia and low concentration CO at high altitude on neurobehavior of drivers. *Highland Med J* 3(1):59–62
2. Ma Y, Zhang R (1999) The effect of hypoxia and stressful task on brain functions. *Chinese Mental Health J* 13(1):48–49
3. Eli I (2016) Subjective feeling and driving safety of highland highway drivers. *Highways Automot Appl* 5:57–60
4. Zhu X, Eli I, Wang Y (2013) Highway traffic safety status and characteristics of accident in plateau area. *Highways Automot Appl* 28(6):78–82
5. Pan X, Song Y, Yang J et al (2009) Area of environmental improvement for highway tunnels import and export based on visual load. *J Tongji Univ (Nat Sci Edn)* 37(6):777–780
6. Du Z, Pan X, Yang J et al (2007) A study on visual shock and traffic safety in highway tunnel import and export. *Chinese J Highway* 05:101–105
7. Hu J (2012) National Natural Science Foundation of China conclusion report: theory and method of calculating safe road conditions
8. Huang F (2014) Experimental study on the improvement of visual illusion of highway drivers based on “black hole effect. Wuhan: Wuhan University of Technology
9. Wang S, Fang J, Han Y (2010) Study on the running speed characteristics of Qinghai-Tibet highway. *China J Highway Transp* 23(1):13–18

10. Feng B, Sun X, Hu Y et al (2016) Effect of air oxygen content on vehicle speed difference in Qinghai-Tibet plateau. *Transp Res* 2(4):45–53
11. Eli S, Li X, Elizati Eli (2015) On-the-spot test and research on the physiological and psychological characteristics of highway drivers on plateau. *China Safety Sci J* 25(12):16–21
12. Khalilikhah M, Heaslip K (2017) Improvement of the performance of animal crossing warning signs. *J Safety Res* 62:1–12
13. ELVIK R (1995) Road design and road equipment. Institute of Transport Economics, Oslo (in Norwegian)
14. Lao Y, Wu Y, Jonathan C et al (2011) Modeling animal–vehicle collisions using diagonal inflated bivariate Poisson regression. *Acc Anal Prevent* 43(1):220–227
15. Liu H (2001) Ergonomic analysis of traffic sign setting at highway intersection. *J Traffic Transp Eng* 1(3):100–103
16. Guo F, Qian S (2000) Human factor engineering. China Machine Press, Beijing
17. Jiang M (2011) Research on prepositive distance calculation model of speed control signs based on driver’s visual recognition and driving characteristics. *Highway* (11):99–103
18. Lu J, Jiang J, Ye H (2011) The method of determining the location of speed limit signs on the road side of common highway. *J Chang’an Univ (Nat Sci Edn)* 31(1):74–78

Analysis on the Road Capacity: Connected and Automated Vehicle Platoons in the Mixed Traffic Flow



Ke Ma and Hao Wang

Abstract The intelligence of connected and automated vehicles (CAV) are not enough to deal with the complex road environment. CAVs based on platoon can reduce external interference and further improve road capacity. Therefore, a mixed traffic flow formula considering ACC, CACC system and leading CAV information interaction based on the CAV platoon is constructed. By analyzing the four main parameters (the size of platoon, ACC headway gain, CACC headway gain and leading CAV communication headway gain), we find the most important platoon parameters that affect the road capacity.

Keywords Connected and automated vehicle · Mixed traffic flow · Capacity analysis · ACC and CACC system

1 Introduction

With the development of urbanization process, urban and highway traffic problem increasingly apparent. It is mainly manifested in the severe traffic congestion, high traffic accident rate and traffic pollution [1–4]. Among them, traffic congestion and safety have attracted much attention from government departments, scholars and engineers. However, if the driving behavior of the participants in traffic activities cannot be effectively improved, problems such as the improvement of traffic operation efficiency [5, 6] and safety [4, 7] will not be fundamentally changed. The

K. Ma · H. Wang (✉)

Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing, China

e-mail: haowang@seu.edu.cn

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

School of Transportation, Southeast University, Nanjing, China

Si Pai Lou, Nanjing 210096, China

K. Ma

e-mail: 220183023@seu.edu.cn

development of Internet of vehicles technology and automated driving technology provides the possibility to solve the above problems. Multi-vehicles collaborative lane-changing behavior [8], speed collaborative optimization [9, 10] and high-speed driving on ramps [11, 12] and intersections have also become new research hotspots based on the development of artificial intelligence and communication technology.

In order to realize the safe and efficient driving of CAVs (Connected and automated vehicles) on the road as soon as possible, foreign scholars made a lot of simulation analysis on the running characteristics of CAVs under V2V (Vehicle to vehicle) environment, involving road capacity, traffic stability and safety. Among them, the most typical applications of V2V communication are Adaptive Cruise Control (ACC) and Cooperative Adaptive Cruise Control (CACC) [13–16], which make CAVs automatically maintain a constant time interval with the preceding vehicle. In both simulation experiments and field experiments, it is generally accepted that continuous CAVs can maintain a smaller headway than traditional MVs (Manual vehicles). Therefore, continuous CAV platoons have become an important factor to improve the highway capacity.

The generation of CAV platoons mainly depend on workshop communication protocol and road side unit communication (V2X) protocol. Queued and continuous CAV platoons can break through the original limitations of sensors in vehicle perception, decision making and control, and ensure vehicle safety by timely acquiring surrounding information to improve driving experience and traffic efficiency. In recent years, scholars at home and abroad have done a lot of research on the traffic efficiency, stability analysis and fuel economy of CAVs traveling in platoons.

So far, most of the work has used the micro simulation or limited theoretical research to draw a conclusion. In order to fill the gaps in the study of lane policies and platoon size, this paper provides a general theoretical formula to illustrate how capability varies with CAV MPR (market penetration rate).

In this paper, capacity refers to the maximum sustainable flow rate of a section given CAV MPR. This is a very important and necessary step in establishing benchmarks to meaningfully understand the impact of micro-characteristics. The microscopic characteristics of CAV (such as platoons length and spacing characteristics) are expressed through a single parameter, the critical headway of CAV, so a clear relationship with the macroscopic capacity is established. We further developed a general formula, including CAV platoon size and headway gain, to determine the distribution and capacity changes of CAVs under a given CAVs communication capability, CAV MPR and headway gain.

2 Switching and Communication

The communication module of CAV system mainly relies on real-time and reliable wireless V2V and V2I. The bandwidth of these two communication methods are insufficient as the CAVs increase within the coverage area. Therefore, in this case,

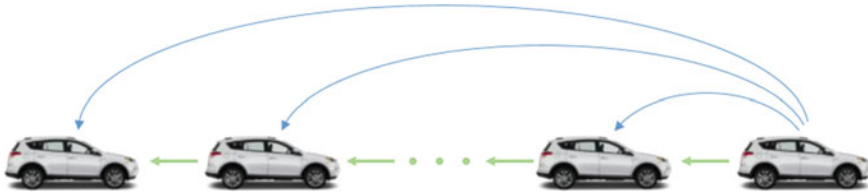


Fig. 1 CAV information flow topologies

short distance wireless technology has advantages. Especially in V2V communication, DSRC (Dedicated Short Range Communications) has been selected as the standard communication protocol. In this paper, we focus on the communication pattern shown in Fig. 1.

We define the information that CAVs can obtain by this communication pattern as follows:

- Information about the status of the preceding CAV including location, speed and acceleration;
- Information from the leading CAV that are beyond sensor range or that are sheltered from view by intermediate CAVs;
- Each CAV that follows an MV automatically become the leader, or the subsequent CAV becomes the leader of the new platoon when the number of CAV reaches the maximum communication scale;

The information transmission mode represented by the green line in Fig. 1 is CACC system, which enables CAVs travel with a low headway. And the operation of CACC system is shown in Fig. 2. CACC can be considered an advanced ACC system as it requires additional activation by DSRC pilot system. In other words, if a CAV follows an MV, the CAV equipped with CACC degrades to AV equipped with ACC. And the blue line represents the leading CAV communicate with the following CAVs in the current platoon.

We assume that both CAVs and MVs travel at a constant free-flow velocity v_{opt} until they reach their critical spacing (We are concerned only with the critical spacing in the presentation of capacity, so the word “critical” is dropped hereafter.) Considering the effective communication range of leading CAV, the length of CAV platoon usually has certain scale limitation. In this paper, CAVs with continuous spatial distribution in the same direction is considered to be a flexible platoon. When the size of the platoon reaches S , the subsequent CAVs are formed into another new platoon. It means that CAV on the road are divided into three categories: leading CAV following MV, leading CAV following CAV and CAV in platoon. We assume that driver cannot identify CAV and MV, therefore, MV use the same car-following model. Then, there are four different headway combinations on the single-lane as shown in Fig. 3, i.e. τ_{m-m} , τ_{a-m} , τ_{c-c} and τ_{cp-cp} . They denote the expected headway for MV following MV (CAV), CAV following MV, CAV following CAV in the same platoon and leading CAV following the last CAV in the preceding CAV.

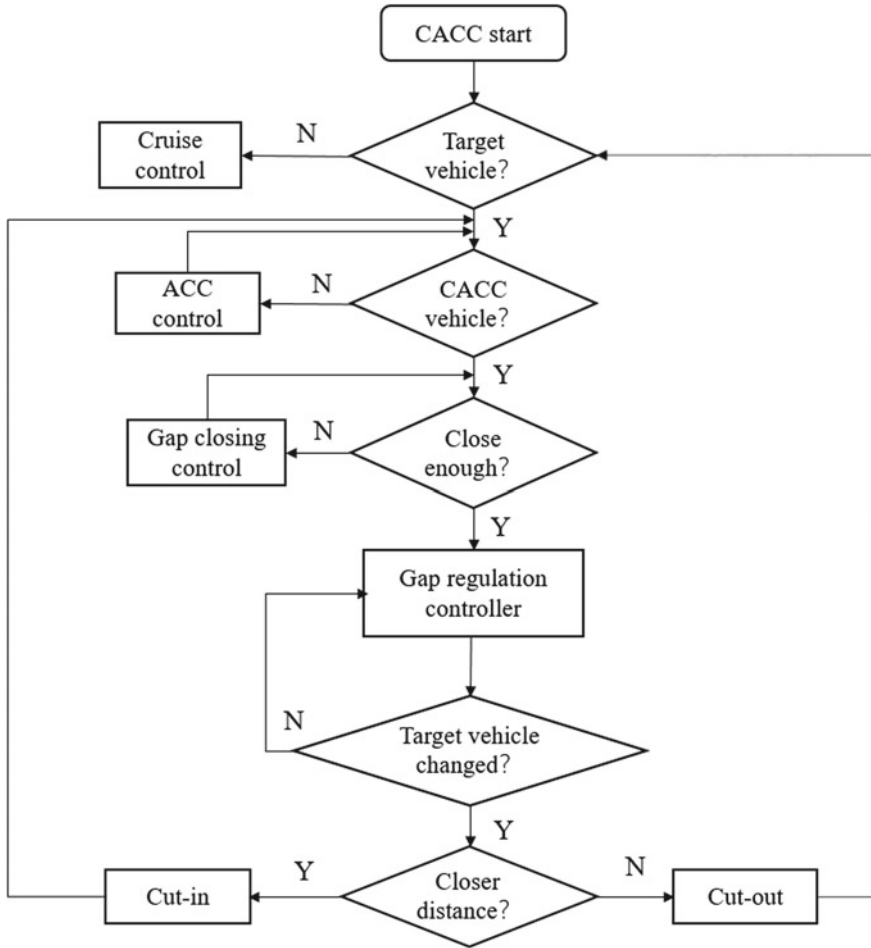


Fig. 2 CACC operation flow chart

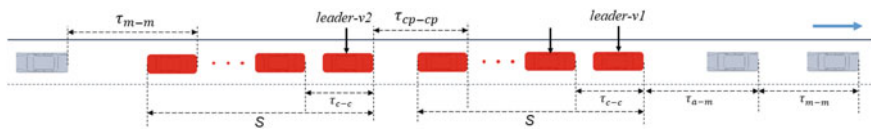


Fig. 3 Illustration of headway in different modes

Obviously, τ_{m-m} is the max headway of the four headway, which is not supported by any on-board sensors or communication equipment. Although some studies suggest that AV with ACC equipment is more sensitive than that of human driving, we believe that AV can operate better based on more accurate detection sensors and

faster decision making modules in the future. Therefore, τ_{a-m} is smaller than τ_{m-m} in this paper. τ_{c-c} can obtain not only the information of the preceding vehicle, but also that of the platoon sent by the leader. So this mode can be seen a more advanced operation than CACC system and τ_{c-c} is the smallest among all the headway. It is worth noting that τ_{cp-cp} also starts CACC system, it is higher than τ_{c-c} due to the lack of preceding platoon communication.

3 Expect Single-Lane Capacity

Initially, we defined capacity as the maximum sustainable flow for given proportions of CAVs and MVs in traffic streams (independent of the CAV MPR, p). And we defined the penetration rate of *leader-v1*, *leader-v2* and other CAV in platoon as p_{v1} , p_{v2} and p_{pl} , respectively. Therefore,

$$p = p_{v1} + p_{v2} + p_{pl}, \quad (1)$$

The penetration rate of MV is

$$p_m = 1 - p, \quad (2)$$

Assuming that all vehicles are mixed randomly and evenly on the road. And we can conclude the penetration rate of *leader-v1* is

$$p_{v1} = p(1 - p), \quad (3)$$

When the number of CAV exceeds, the next CAV become the new leader of another platoon, and so on. Then, the penetration rate of *leader-v2* is

$$p_{v2} = \frac{p^{s+1}(1 - p)}{1 - p^s}, \quad (4)$$

The penetration rate of CAVs in platoon is

$$p_{pl} = \frac{p^2 \times (1 - p^{s-1})}{1 - p^s}, \quad (5)$$

We can assume that there are m CAVs and n MVs on the road. Therefore, the average headway of all vehicles on the single-lane is

$$\tilde{h} = p_{v1} * \tau_{a-m} + p_{v2} * \tau_{cp-cp} + p_{pl} * \tau_{c-c} + p_m * \tau_{m-m}, \quad (6)$$

Therefore, road capacity for the mixed traffic flow is

$$C_m = \frac{1}{\tilde{h}} = \frac{1}{p_{v1} * \tau_{a-m} + p_{v2} * \tau_{cp-cp} + p_{pl} * \tau_{c-c} + p_m * \tau_{m-m}}, \quad (7)$$

To simplify this formula, we use a matrix to represent the headway gain. Rewrite Eq. 6,

$$\tilde{h} = \mathbf{P}^T \cdot \boldsymbol{\lambda} \cdot \tau_{m-m}, \quad (8)$$

where, \mathbf{P} represents vehicle penetration rate vector,

$$\mathbf{P} = \begin{bmatrix} p_{v1} \\ p_{v2} \\ p_{pl} \\ p_m \end{bmatrix}, \quad (9)$$

Similarly, $\boldsymbol{\lambda}$ represents the headway gain of CAVs to MVs under different modes,

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_a \\ k_c \lambda_a \\ k_p k_c \lambda_a \\ 1 \end{bmatrix}, \quad (10)$$

where, λ_a , k_c and k_p represents the headway gain of ACC system to MVs, the gain of CACC to ACC system and the gain of inter-platoon. Assuming that the capacity without CAVs, C_0 , on the road is constant, so the capacity gain of C_m to C_0 is

$$\frac{C_m}{C_0} = \frac{1}{\mathbf{P}^T \cdot \boldsymbol{\lambda}}, \quad (11)$$

4 Simulation and Conclusion

We performed experiments on different parameters to compare the effects of different parameters on the road capacity based on the Eq. (11). When other parameters are constant, S , λ_a , k_c and k_p are respectively analyzed under different CAV MPR conditions.

As shown in Figs. 4 and 5, road capacity increases significantly with the increase of CAV MPR even though under different parameters. Moreover, as the increasing of S , λ_a , k_c and k_p , the mixed capacity gain has increased. Through a series of simulations and calculations, we can draw some important conclusions:

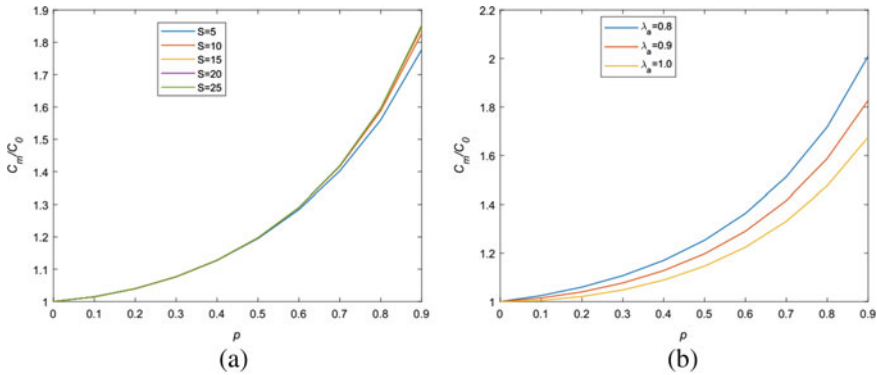


Fig. 4 **a** Capacity gain changes with various p (S from 5 to 10; $\lambda_a = 0.9$, $k_c = 0.8$ and $k_p = 0.7$)
b Capacity gain changes with various p ($S = 15$; λ_a from 0.8 to 1.0, $k_c = 0.8$ and $k_p = 0.7$)

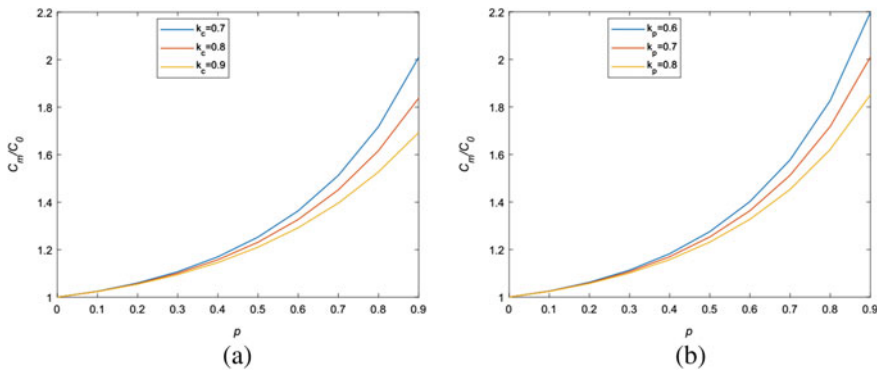


Fig. 5 **a** Capacity gain changes with various p ($S = 15$; $\lambda_a = 0.9$, k_c from 0.7 to 0.9 and $k_p = 0.7$)
b Capacity gain changes with various p ($S = 15$; $\lambda_a = 0.9$, $k_c = 0.8$ and k_p from 0.6 to 0.8)

- (1) The change of capacity gain is nonlinear with the change of CAV MPR. This means that when CAV MPR is low, road capacity will not be improved significantly. However, with the further increase of the number of CAV, capacity will be rapidly improved.
- (2) The effect of the size of the CAV platoon is obviously minimal. Although the development of communication range and speed is very rapid, its impact in improving the road capacity is not significant.
- (3) Obviously, improving inter-platoon communication technology can maximize the mixed capacity gain. That is because this part has the most CAVs and a further reduction in the inter-platoon CAVs headway can yield the most significant benefits.

The rapid development of autonomous vehicles in intelligent environments has attracted the attention of many researchers. In order to ensure the safe and stable

operation of CAVs, researchers have presented important models and tests in many aspects. In this paper, the influence of multiple parameters on capacity is modeled, and the parameters that can improve capacity more effectively are further pointed out.

Acknowledgements This research was sponsored by the National Key Research and Development Program of China (No. 2018YFB1600900).

References

1. Vander Laan Z, Sadabadi KF (2017) Operational performance of a congested corridor with lanes dedicated to autonomous vehicle traffic. In *J Transp Technol* 6(1):42–52
2. Ma K, Wang H (2019) A cellular automaton model considering the exclusive lanes of autonomous vehicles on expressway. In: *CICTP 2019: tTransportation in China—connecting the world—proceedings of the 19th COTA international conference of transportation professionals*
3. Wang, Meng (2018) Infrastructure assisted adaptive driving to stabilise heterogeneous vehicle strings. *Transp Res Part C Emerging Technol* 91:276–295
4. Ye et al (2017) Evaluation of the impacts of cooperative adaptive cruise control on reducing rear-end collision risks on freeways. *Accident Analysis Prevention*
5. Hao, Wang et al (2018) Stability of CACC-manual heterogeneous vehicular flow with partial CACC performance degrading. *Transp B Transp Dyn* 1–26
6. Chen, D et al (2017) Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles. *Transp Res* 196–221
7. Zhu W-X, Li-Dong Z (2018) A new car-following model for autonomous vehicles flow with mean expected velocity field. *Phys A Statist Mech Appl* 492:2154–2165
8. Liu FY, Wang H, Dong C, Chen Q (2020) a car-following data collecting method based on binocular stereo vision. *IEEE Access* 8:25350–25363. <https://doi.org/10.1109/ACCESS.2020.2965833>
9. Sun et al (2018) An extended car-following model under V2V communication environment and its delayed-feedback control. *Phys A Statist Mech Appl*
10. Stern RE et al (2018) Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. *Transp Res Part C Emerging Technol* 89:205–221
11. Goodall, NJ, Park B, Smith BL (2014) Microscopic estimation of arterial vehicle positions in a low-penetration-rate connected vehicle environment. *J Transp Eng* 140(10):04014047.1–04014047.9
12. Milanes V, Shladover SE (2014) Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transp Res* 48:285–300
13. Ma, K e . “Influence of Exclusive Lanes for Connected and Autonomous Vehicles on Freeway Traffic Flow.” *IEEE Access* 7(2019).
14. Qin Y, Hao W, Bin R (2018) Stability analysis of connected and automated vehicles to reduce fuel consumption and emissions. *J Transp Eng Part A Syst* 144(11):04018068
15. Dong C et al (2020) Economic comparison between vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) at freeway on-ramps based on microscopic simulations. *IET Intell Transp Syst* 13(11):1726–1735
16. Ge, Ji et al (2018) Experimental validation of connected automated vehicle design among human-driven vehicles. *Transp Res* 91:335–352

Real-time Trajectory Prediction of Unmanned Aircraft Vehicles Based on Gated Recurrent Unit



Rong Tang, Zhao Yang, Jiahuan Lu, Hao Liu, and Honghai Zhang

Abstract In view of the increasing number of unmanned aircraft vehicles (UAVs) in low altitude airspace, which leads to the increasing difficulty of surveillance, a real-time trajectory prediction method based on the gated recurrent unit (GRU) is proposed. The parameters of the GRU model are determined including the number of hidden layers, the cell sizes of a single hidden layer and the time step. Historical trajectory data are used to train and test the model. The result shows that the mean absolute errors (MAE) of altitude, longitude, and latitude predicted by the GRU model are 0.86 m, 2.85 m, and 3.65 m, respectively. The prediction results of the GRU model are compared with the long and short-term memory (LSTM), the support vector regression (SVR) and the autoregressive integrated moving average model (ARIMA). It is found that the GRU and LSTM outperform the two other models. The LSTM prediction accuracy is slightly higher than that of the GRU model. However, the space complexity of the LSTM model is much larger than that of the GRU model, and its training time is much longer.

Keywords Unmanned aerial vehicle · Trajectory prediction · Gated recurrent unit · Machine learning · Data mining

1 Introduction

With the rapid development of the unmanned aircraft vehicle (UAV) industry, there are more and more small and medium-sized consumer UAVs in the low-altitude airspace of urban areas. Incidents that threaten the safety of low-altitude airspace by UAV activities frequently occur. Most consumer-grade UAVs use a four-rotor design structure, which have flexible movements. Real-time and accurate trajectory prediction will provide technical support for identification of conflicts and dangerous intentions in low-altitude airspace.

R. Tang · Z. Yang (✉) · J. Lu · H. Liu · H. Zhang
National Key Laboratory of Air Traffic Flow Management, College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Jiangjun Road No. 29, Nanjing 211106, China
e-mail: seu_yolanda@126.com

At present, research on the UAV trajectory prediction can be roughly divided into three categories: dynamics-based prediction models, machine learning-based prediction models and hybrid prediction models. The dynamics-based prediction models are conducted based on the behavioral analysis of the UAVs by establishing high-order nonlinear motion equations to investigate the UAV trajectory characteristics [1–3], combined with the UAV specific motion pattern recognition [4–6]. For the rotary-wing UAV, composite motion models of translation and rotation are established to predict the future movements [7]. The machine learning methods are applied to predict the trajectory of UAVs by learning from historical trajectory data. Some commonly used methods include gray theory, Markov [8, 9] and neural networks [10, 11]. The hybrid prediction method is mainly based on the optimal theory, such as Kalman filtering and other algorithms [12, 13].

Although some studies have been conducted in the aspect of trajectory prediction, most of them are focused on fixed-wing aircraft or commercial aircraft [14–16]. Due to the flexible and changeable motions of slow-speed UAVs such as rotors, the kinematics modeling procedure is very complex. Besides, many assumptions are required to establish the motion equation. With the advance of machine learning techniques, the prediction accuracy of UAVs can be improved. Research on the use of machine learning to achieve trajectory prediction often requires a large number of features and historic samples [17–19]. These conditions are difficult to be satisfied in the environment of short-time trajectory prediction. In this paper, a real-time trajectory prediction model for slow-speed UAVs is established. The input data is relatively simple, convenient for real-time collection, and the prediction time is short, which meets the time requirement of fast prediction.

2 Data Collection and Preparation

For the real-time prediction of the UAV trajectory, this paper collected a total of 800 historical trajectory data of a small quadrotor UAV, including time, altitude, longitude, latitude and ground speed, at 5 s intervals, resulting in a total of 59,900 samples. According to the ultimate speed of the experimental UAV, the five-second motion limit value of its three dimensions is obtained. Data of the 5-second motion range of any dimension in the sample that exceeds the limit value is removed. The removed position and the missing data position are filled with the average value of the front and back positions.

The historical trajectory data of each flight of the UAV is regarded as a subsample. Each subsample is processed separately, and then combined to form a total sample. The training set, the verification set and the test set are divided. The specific steps are listed as follows:

Step 1: Determine the input and output indicators. The input indicators are altitude, vertical speed, longitude, latitude and horizontal speed, and the output indicators are altitude, latitude and longitude.

Step 2: Extract the required data in Step 1 from the historical flight record data. The data are then normalized using the min-max methods.

Step 3: Set the initial values of the GRU network parameters, including the time step, the number of neural network layers, and the number of neurons in a single layer.

Step 4: Divide the trajectory data of a UAV into multiple pieces of data according to the data recording time label into sub-samples. Each piece of data in the sub-sample contains two parts, including the inputs and outputs. The outputs are the index values that need to be predicted. The inputs are the numerical sequence of indicators with determined timestep. The dimension of the input part is [number of subsample data, determined timestep, number of input indicators], and the dimension of the output part is [number of subsample data, number of output indicators].

Step 5: Combine the input and output of multiple subsamples processed in Step 4 according to the first dimension [the number of subsample data] to form the input and output of the total sample. The ratio of training dataset to test dataset is set to be 3:1.

3 Methodology

In this section, a gated recurrent unit (GRU) model is constructed for intention identification and trajectory prediction of quadrotors [20]. The procedure is presented as follows.

Similar to long short-term memory (LSTM), GRU is also used to solve the long-term memory problem existing in RNN and the gradient problem in backpropagation. However, different with the structure of forgetting gate, input gate, cell gate and output gate in LSTM, the GRU network has reset gate and update gate. The function of update gate is similar with the combination of the forgotten gate and the input gate in the LSTM, which is responsible for processing the memory and the current input information. The principle of a GRU unit is shown in Fig. 1.

As shown in Fig. 1, the input of a single GRU neuron contains two parts of information, the current state information x_t of the UAV and the state feature h_{t-1} retained at the previous moment. The output also contains the state feature h_t , which is the same as the output y_t . The specific principle of a single GRU neuron is listed as follows:

First, use the sigmoid function to activate the input information x_t and hidden state h_{t-1} so as to obtain the current reset gate state r_t . The formula is:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (1)$$

where x_t is the input information at the current moment; h_{t-1} is the hidden state at the previous moment; W_r is the weight corresponding to the input information x_t and the hidden state h_{t-1} , which is updated continuously with training. Similarly, the current state z_t of the update gate can be calculated by the following formula:

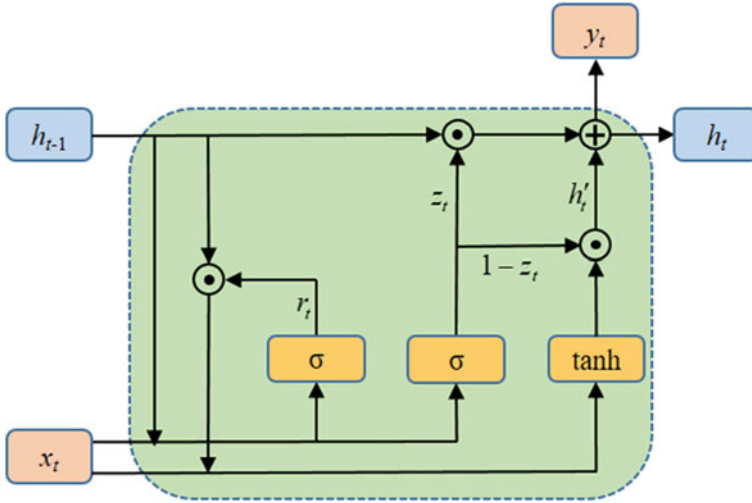


Fig. 1 Schematic diagram of a single GRU cell

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \tag{2}$$

where W_z is the weight corresponding to x_t and h_{t-1} in the update gate

Second, the reset gate state r_t , the hidden state h_{t-1} , and the input x_t are combined to calculate the memory information h'_t at the current moment. The formula is:

$$h'_t = \tanh(W_t \cdot [r_t \odot h_{t-1}, x_t]) \tag{3}$$

where W_t is the weight corresponding to x_t and reset gate state r_t multiplied by hidden state h_{t-1} . When r_t is 0, the memory information at the current time is determined only by the input x_t .

Third, the update gate state z_t is applied to forget and retain the information contained in h_{t-1} and h'_t respectively. The formula is:

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h'_t \tag{4}$$

where h_t is the output of neurons at the current moment and also the hidden state passed to the next moment.

In order to prevent over-fitting due to insufficient samples, L2 regularization is added as the loss function on the mean square error. The specific formula is as follows:

$$Loss = \frac{1}{n} \sum_{i=1}^n (y(x)_i - y_i)^2 + \lambda \|W\|_2^2 \tag{5}$$

where y_i is the real information of the UAV motion at time i , W is the weight set required to calculate the predicted information y from the input x , and $y(x)_i$ is the predicted value of the UAV motion information at time i .

The mean absolute error (MAE) is chosen as the loss evaluation function for trajectory prediction of quadrotors, which can be calculated as follow:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y'_i - y_i| \quad (6)$$

where y'_i are the predicted value in i -th trajectory point, y_i are the real value in i -th trajectory point.

4 Data Analysis Results

4.1 Determination of hyperparameters for the GRU method

In order to obtain the appropriate network structure, the control variable method is used to determine the number of hidden layers h in the network, the cell size c in a single layer, and the time step t of the input information.

First, the initial value of t is set to be 5, c is set from 5 to 30 with interval 5 (5/30/5), and h is set from 1 to 9 with interval 2 (1/9/2). The MAEs are calculated. The results are shown in Fig. 2. It can be known that when h is 3, the MAE can reach a relatively low value under different values of c . When h increases above 5, the MAE starts to increase. It is due to the reason that the demand for the number of input indicators and the number of training samples are increased when the depth of the network is increased. For the real-time trajectory prediction of UAV, the number of input features and the training samples are few, and the training result will be over-fitting, which leads to the phenomenon that MAE increases rapidly when the network depth is large. Although 10% of the neuron output are randomly dropped out and L2 regularization are used to reduce the over-fitting phenomenon of the network, the MAE is still higher than that of the network when h is 3.

Since the MAE when h equals 3 is lower than that of other parameter settings, the MAEs are calculated when c is set from 5 to 50 with interval 5 (5/30/5), and h is set from 2 to 4 with interval 1 (2/4/1) when t is 5. In order to identify the effect of a larger value of c on the MAE, the experimental result when c is set to 50 is presented, as shown in Fig. 3. It can be found that the MAE changes slightly when h is set to 2, 3 and 4 with different values of c . The MAE is slightly increased compared to the other parameter settings when c is set to 50. Together with Fig. 2, it can be found that h is a more important factor affecting the MAE as compared with c . Finally, h and c are set to be 3 and 5, respectively.

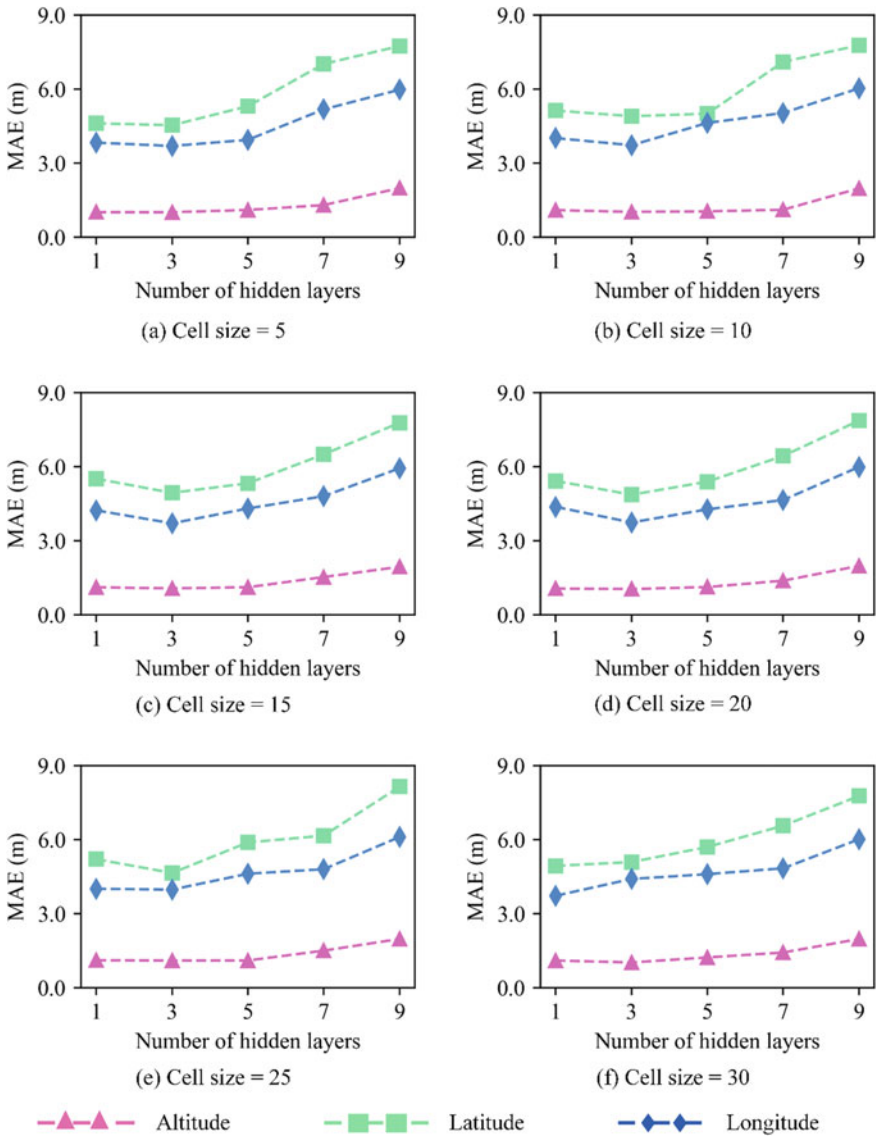


Fig. 2 MAEs with respect to the number of hidden layers

As shown in Fig. 4a, the MAE of the GRU model decreases by about 1m as t increases from 10 to 50 for both longitude and latitude, while the MAE for altitude is almost invariant. When t is 40, the MAE of the GRU model reaches the local minimum value, so 40 is selected as the set value of t of the GRU model. Further, the parameters of the GRU network (without t) are applied in the LSTM model, and the change of MAE is tested as t increases. The result is shown in Fig. 4b. The changes

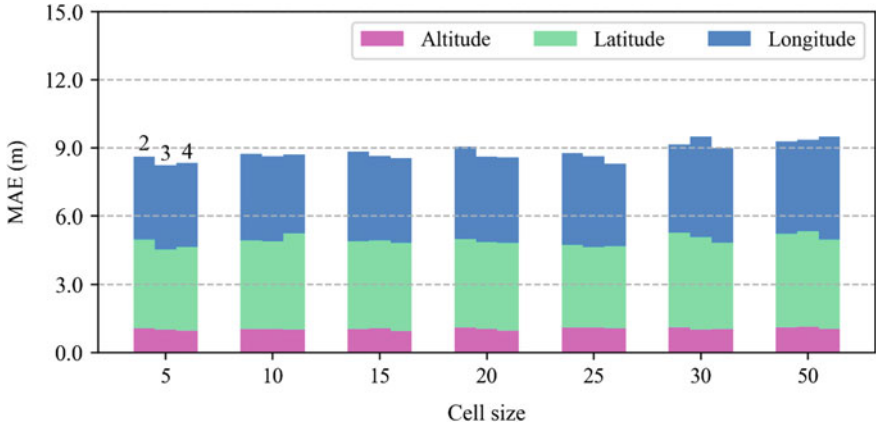


Fig. 3 Accumulated errors with different number of single-layer neurons and hidden layers

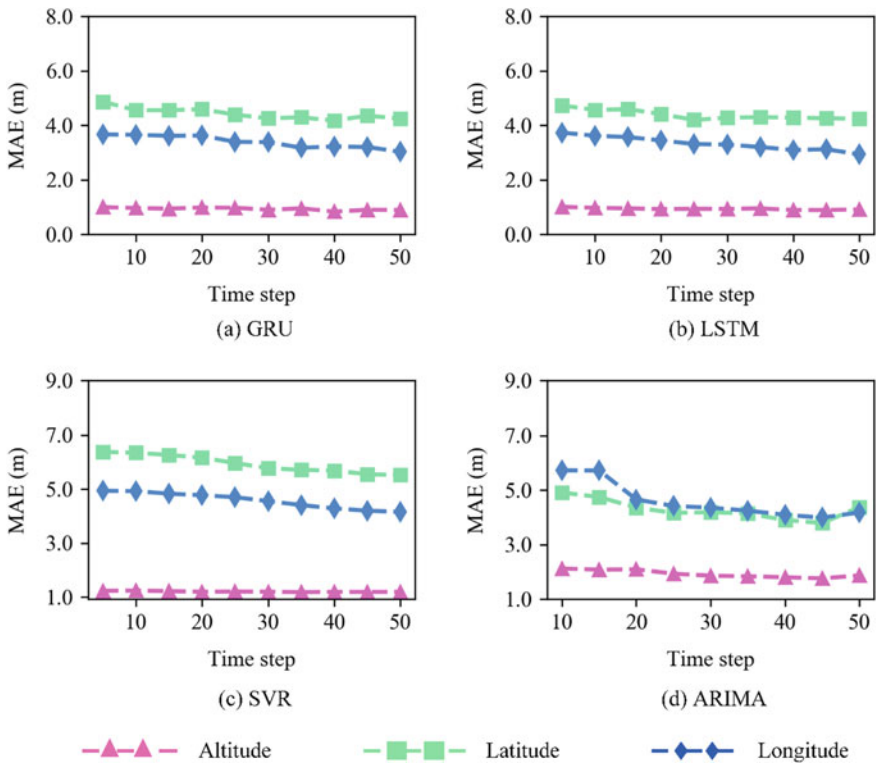


Fig. 4 Performance comparison of the four models with different time steps

Table 1 Comparison of MAEs when the four models take the optimal timestep (m)

	GRU	LSTM	SVR	ARIMA
The optimal timestep	40	50	50	45
Altitude	0.86	0.80	1.20	1.80
Latitude	3.65	3.49	5.53	4.10
Longitude	2.85	2.73	4.16	3.91

of MAE of the SVR and ARIMA models are shown in Fig. 4c and d when t is set to different values.

According to Fig. 4, The minimum MAEs and corresponding time steps of the four models are shown in Table 1. The MAE difference between of the GRU and LSTM models are 0.06m, 0.16m, and 0.12m for altitude, latitude and longitude, respectively, indicating that the performance of the GRU model is general comparable to that of LSTM with determined parameters in the network. In contrast, the MAE for GRU are 28%, 11% and 27% lower than that of the SVR and ARIMA models in altitude, latitude and longitude, respectively. The results also show that the MAE of latitude is larger than that of longitude, which is due to the reason that the data has a larger variation in the latitude direction. The MAE value is larger for latitude after denormalization.

Figure 5 illustrates the comparison of actual position and prediction results for a randomly selected trajectory as functions of time. The x -axis represents time and the y -axis represents latitude in Fig. 5a and b, longitude in Fig. 5c and d and altitude in Fig. 5e and f. The grey line from time 0 to 250 sec represents historic positions that used as input values, and the grey line from time 250–1500 sec represents true positions that used as the baseline for comparison. The red and purple lines represent the prediction results of LSTM and GRU methods, respectively. It can be found that the LSTM and GRU methods are generally comparable.

Figure 6 demonstrates the comparison of training time and training loss for GRU and LSTM methods. The batch size is set to be 100. The 30 batches of training information are recorded. The convergence of training loss and distribution of training time of the GRU and LSTM models are shown in Fig. 6, where the training loss of the GRU model is marked with pink, the training time is marked with cyan, and the corresponding gray part is the result of the LSTM model. It shows that the convergence speeds of the two models are basically the same. The two models have reached a convergent and stable state after training for 5 seconds. Since the time consumption of predicting is far less than training, this method can meet the real-time trajectory UAV prediction.

It can also be found from Fig. 6 that as the number of training increases, the training time of GRU will be shorter than that of LSTM. The reason for this phenomenon is the difference in complexity of the two models. When the GRU and LSTM models take the same timestep, their input and output dimensions and hyperparameters are the same, and the number of iterations of network training is also the same, which means that the time complexity is the same. Therefore, the difference in complexity

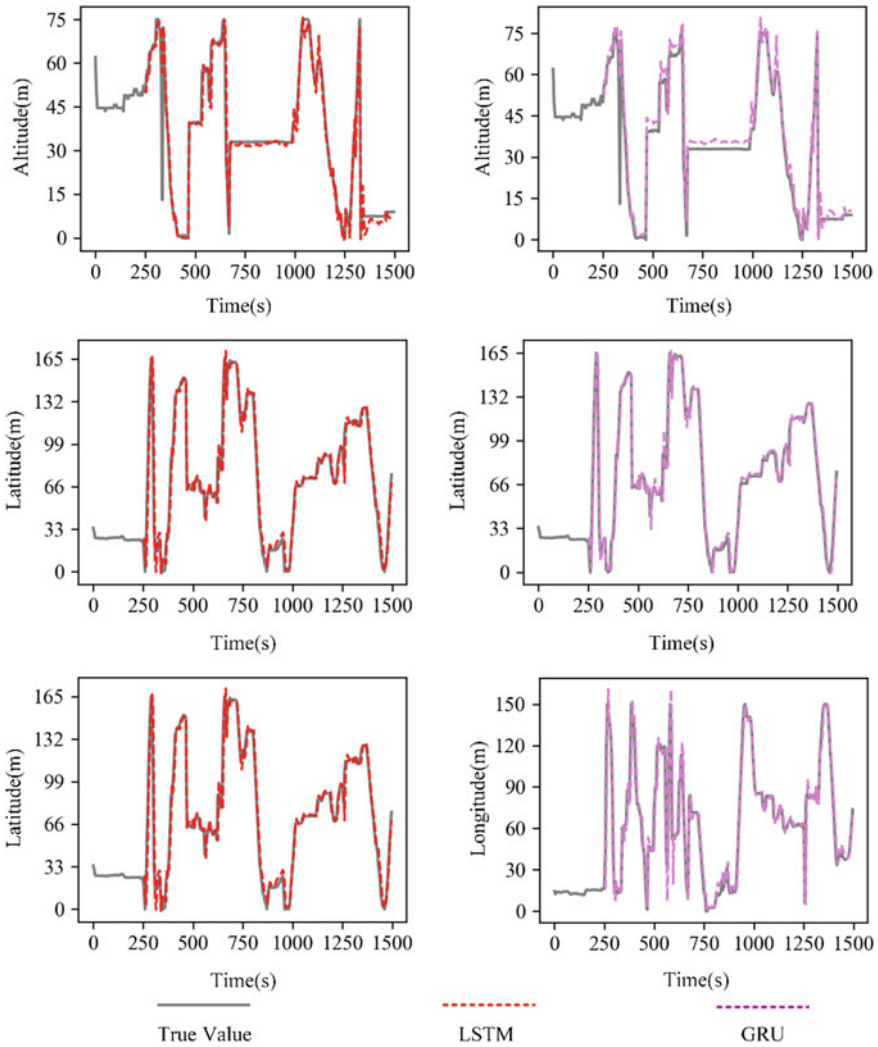


Fig. 5 Comparison of true position and prediction results

between the two models is mainly concentrated in space complexity. Since there is no encoding and decoding process in this work, the parameters of the network are mainly concentrated in the hidden layers. Thus, the problem is simplified to compare the parameter amount of the hidden layer. The number of the biases is far less than the weights, so biases are not considered.

Table 2 illustrates the comparison of space complexity between GRU and LSTM models. It can be seen from the formula in Table 2 that single-layer GRU has one less gate than single-layer LSTM. The weight parameters of GRU are W_r , W_z and W_l , while the weight parameters of LSTM are W_f , W_i , W_c and W_o , respectively. When

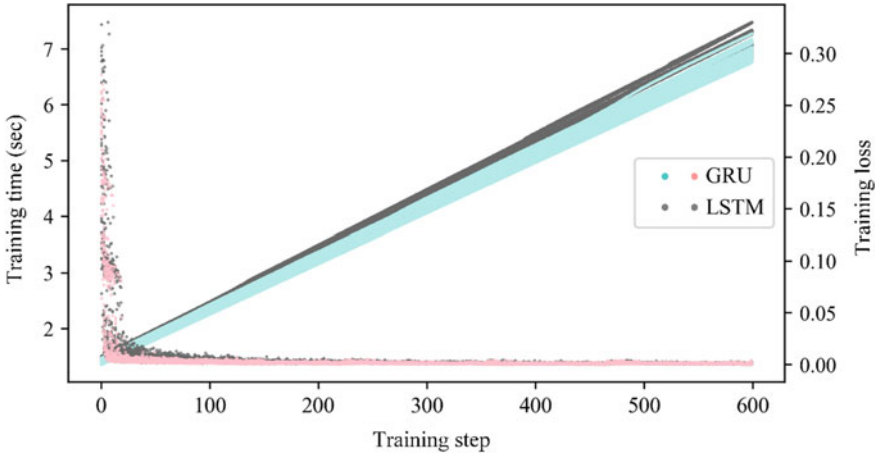


Fig. 6 Comparison of training time and training loss for GRU and LSTM

Table 2. Comparison of space complexity between GRU and LSTM

Models	GRU	LSTM
Equations	$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$ $z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$ $h'_t = \tanh(W_t \cdot [r_t \odot h_{t-1}, x_t])$ $h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h'_t$	$f_t = \sigma(W_f \cdot [h_{t-1}, x_t])$ $i_t = \sigma(W_i \cdot [h_{t-1}, x_t])$ $C'_t = \tanh(W_c \cdot [h_{t-1}, x_t])$ $C_t = f_t \odot C_{t-1} + i_t \odot C'_t$ $o_t = \sigma(W_o \cdot [h_{t-1}, x_t])$ $h_t = o_t \odot \tanh(C_t)$
Weights	W_r, W_z, W_t	W_f, W_i, W_c, W_o

the two hidden layer structures (the number of layers and the number of neurons) are the same, the total number of parameters of a single-layer GRU is nearly 25% less than that of a single-layer LSTM, regardless of the number of network layers and training times. Therefore, as shown in Fig. 6, GRU is much faster than LSTM in data training. As the training time increases, the absolute time difference between the predicted value and the true value will also increase. The finding is consistent with that of previous studies. Through comparative experiments, Chung et al. also confirmed that GRU tends to converge faster than LSTM, but at the same time the authors pointed out that this does not mean that GRU is better than LSTM. It is necessary to choose a suitable model according to specific experimental tasks [21]. For real-time prediction of UAV trajectory within a certain number of training times (before convergence), the error is inversely proportional with time-consuming. The results indicate that GRU model is more suitable for real-time UAV trajectory prediction, as long as the MAE meets the requirement of forecasting accuracy.

5 Conclusions

This paper establishes a GRU model for real-time prediction of UAV trajectories based on historical trajectory data. The average absolute error of altitude prediction is 0.86 m, and the average absolute error of latitude and longitude prediction are 2.85 m and 3.65 m, respectively. The convergence speed of the model is extremely fast, which can meet the time requirement of real-time prediction of UAV trajectory. Results of the GRU model is compared with those of the other methods including LSTM, SVR and ARIMA. It is found that the prediction accuracy of the GRU model is very close to that of the LSTM model, and outperforms the traditional SVR and ARIMA models. Besides, the GRU model has a faster training speed and is more suitable for the task of real-time trajectory prediction. Further analysis can be conducted by combining UAV historical trajectory data, UAV intention recognition and UAV flight dynamics model to further improve the accuracy of UAV real-time trajectory prediction.

6 Funding Statement

This research was sponsored by National Natural Science Foundation of China (52172328, 62076126, 52002179), the Fundamental Research Funds for the Central Universities of China (NS2020046), and Science and Technology Innovation Project for College Students (2020CX00760, 2020CX00753, kfjj20200713, 202110287093Z, 2021CX020003).

References

1. Alexis K, Nikolakopoulos G, Tzes A (2014) On trajectory tracking model predictive control of an unmanned quadrotor helicopter subject to aerodynamic disturbances. *Asian J Control* 16(1):209–224
2. Sun X, Liu Y, Yao W et al (2015) Triple-stage path prediction algorithm for real-time mission planning of multi-UAV. *Electron Lett* 51(19):1490–1492
3. Ishihara AK, Jung J, Rios J (2016) Rapid trajectory prediction for a fixed-wing UAS in a uniform wind field with specified arrival times. AITO 2016: 16th AIAA aviation technology, integration, and operations conference. AIAA, Washington, D.C., pp 1–19
4. Jeung H, Liu Q, Shen HT et al (2008) A hybrid prediction model for moving objects. In: ICDE 2008: IEEE 24th international conference on data engineering. IEEE, Cancun, 70–79
5. Maeder U, Morari M, Baumgartner TI (2011) Trajectory prediction for light aircraft. *J Guidance Control Dyn* 34(4):1112–1119
6. Lin CE, Lai Y (2018) UAV path prediction for CD&R to manned aircraft in a confined airspace for cooperative mission. *Int J Aerospace Eng* 1:1–9
7. Corbetta M, Banerjee P, Okolo WA et al (2019) Real-time UAV trajectory prediction for safety monitoring in low-altitude airspace. AIAA Aviation Forum. AIAA, Dallas, pp 1–17
8. Ke H, F. Liusi F, Chen Y (2010) A real-time vision prediction algorithm for aviation track of unmanned air vehicle. In: CCC 2010: proceedings of the 29th Chinese control conference. IEEE, Beijing, 3098–3102

9. Qiao S, Shen D, Wang X et al (2015) A self-Adaptive parameter selection trajectory prediction approach via hidden markov models. *IEEE Trans Intell Transp Syst* 16(1):284–296
10. Xue M (2017) UAV trajectory modeling using neural networks. AITO 2017: 17th AIAA aviation technology, integration, and operations conference. AIAA, Denver, pp 1–10
11. Xiao K, Zhao J, He Y et al (2019) Trajectory prediction of UAV in smart city using recurrent neural networks. In: ICC 2019: 2019 IEEE international conference on communications. IEEE, Shanghai, 1–6
12. Prevost CG, Desbiens A, Gagnon E (2007) Extended kalman filter for state estimation and trajectory prediction of a moving object detected by an unmanned aerial vehicle. In: ACC, 2007 proceedings of the 2007 American control conference. ACC, New York, 1805–1810
13. Zhang X, Fan C, Fang J et al (2018) Tracking prediction to avoid obstacle path of agricultural unmanned aerial vehicle based on particle filter. *J Syst Control Eng* 232(4):408–416
14. Stastny T, Dash A, Siegwart R. Nonlinear MPC for fixed-wing UAV trajectory tracking: implementation and flight experiments. In: AIAA guidance, navigation, and control conference, 1–14
15. Gavilan F, Vazquez R, Esteban S (2015) Trajectory tracking for fixed-wing UAV using model predictive control and adaptive backstepping. In: IFAC-PapersOnLine, 48–9, 132–137.
16. Guan X, Lv R, Sun L et al (2016) A study of 4D trajectory prediction based on machine deep learning. In: WCICA 2016: 12th world congress on intelligent control and automation. WCICA, Guilin, 24–27
17. Shen M, Xie Y, Xue X et al (2009) Study on track prediction of air target based on grey system model. *Kybernetes* 38(3/4):468–473
18. Valsamis A, Tserpes K, Zissis D et al (2016) Employing traditional machine learning algorithms for big data streams analysis: the case of object trajectory prediction. *J Syst Softw* 127:249–257
19. Zhao Z, Zeng W, Quan Z et al (2019) Aircraft trajectory prediction using deep long short-term memory networks. In: CICTP 2019: 19th COTA international conference of transportation professionals. CICTP, Nanjing, 124–135
20. Cho K, Merriënboer BV, Gulcehre C (2014) Learning Phrase representations using RNN encoder–decoder for statistical machine translation. arXiv preprint arXiv: 1406.1078v3
21. Chung J, Gulcehre C, Cho K (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv: 1412.3555v1

The Capacity Model of Urban Multimode Transportation Network Based on Combined Travel Mode



Guozhu Cheng and Chuanmiao Zhou

Abstract In order to research the carrying capacity of the urban multimode transportation network to for the number of trips, this paper constructs the corresponding super network and travel cost functions based on combined travel modes. For a given current O-D pattern, a bi-level programming model is formulated to estimate the reserve capacity of the multimode transportation network, the upper-level model aims at maximizing the amount of travel, the lower-level model describes the traveler's behavior, conforms to the stochastic equilibrium distribution. Finally, considering the limitation of transfer time and level of service, a heuristic method based on incremental assignment technique is proposed to solve this model.

Keywords Super network · Combined travel mode · Reserve capacity · Bi-level programming

1 Introduction

Capacity modeling of a transportation network ought to be an essential part of any urban and transportation systems planning, which reflects the bearing capacity of a given transportation network, providing quantitative data for traffic planning, traffic management and control. According to existing research, several main design methods are summarized as following.

- (1) Time–space consumption method: by calculating the time–space resource consumption of the transportation network to obtain the network capacity, scholars can use their engineering experience to revise the road capacity and service time, which made the result more in line with the actual situation [1, 2].
- (2) Cut set method [3, 4]: applying the max-flow min-cut set theorem in graph theory to solve. The network capacity calculated by this method is defined as

G. Cheng (✉) · C. Zhou

School of Traffic and Transportation, Northeast Forestry University, Harbin 150040, China

e-mail: guozuhucheng@126.com

the passing capacity of the smallest section separating all cut sets between OD nodes.

- (3) Nonlinear bi-level programming: solve the network capacity by constructing a non-linear bi-level programming model. The upper-level model is objective of the planner, in order to maximize the service volume, while taking into account some constraints: level of service [5], toll road [6], environmental pollution [7].

In addition, changing the decision-making objective can solve some network optimization problems, in the upper-level model, maximize the capacity while minimizing performance index of the network to optimize traffic signal timings [8]. The biobjective programming model can also be used to solve the Continuous Network Design (CND) [9] problem and Discrete Network Design (DND) problem [10] problems. The lower-level model is the description of traveler's behavior. According to the problem, the combined traffic distribution-assignment model is usually selected. Some studies have explored the impact of demand distribution [11] and traveler's behavior changes on the network. The lower-level model reflects the traveler with respect to the difference in the quality of path information [12] and path preference [13], the network capacity doesn't increase monotonically as the quality of road information available to travelers improves.

Most previous studies focused on describing traveler's behavior from the perspective of a single travel mode or mode split, doing not take into account the transfer of multimodal transportation modes, and there are insufficient descriptions of traveler's behavior. Based on this, this paper transforms the urban multimode transportation network in a super network, and evaluates the bearing capacity of the urban multimode transportation network in the perspective of combined travel modes. The remainder of this paper is structured as follows. In Sect. 2, the corresponding super network is constructed. In Sect. 3, the bi-level programming model and the solving algorithm for solving the capacity are given. An example of the model application and the conclusion are shown in Sects. 4 and 5.

2 Network Structure

The sub-networks corresponding to each travel mode in the city are combined into a multimode transportation network, and each network has its own independent structure. Different networks are connected through transfer activities, and the transfer behaviors have a certain displacement in space, which will generate travel costs in a broad sense. It can be regarded as a virtual link connecting all layers of the network, thus forming an interconnected multilayer network, as shown in Fig. 1.

Based on the super network and expansion theory, the multilayer transportation network shown in Fig. 1 is transformed into the super network structure shown in Fig. 2. Each link in Fig. 2, including the actual link and the transfer link, has attributes such as length and travel cost. let $G = (N, A)$ be the network defined by a set N of nodes and a set of A of directed links a .

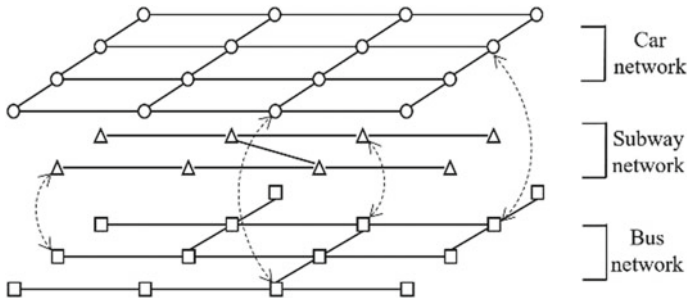


Fig. 1 Multilayer transportation network

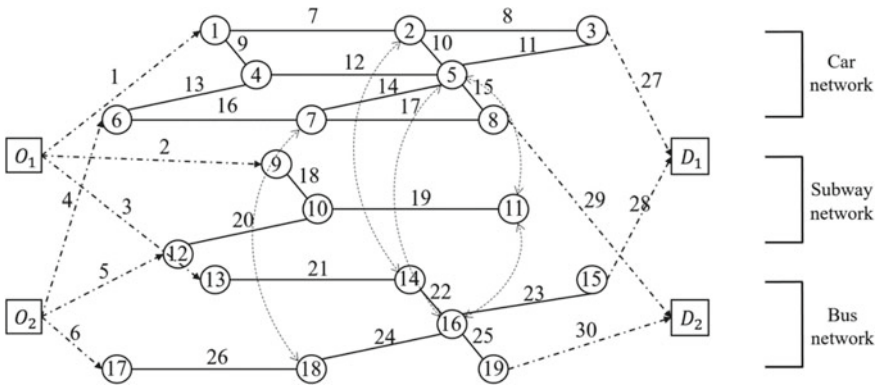


Fig. 2 Super network

2.1 Travel Costs

The process from origin to destination of the traveler can be partitioned as: entering the network, choose a way to travel and transfer to another, at the end, leave the network. In this paper, we don't think about cycling, the way to go in and out of the network and transfer is walking.

When traveler make choices about their routes, the main considerations are travel time, economic cost, comfort, and the number of transfers, among them, travel time is the most important. In the constructed super network, total time consumed is composed of travel time and transfer time.

The total travel cost includes time cost, economic cost and comfort cost. The travel cost function converted into generalized cost is shown in Formula (1):

$$c_w = t_a + \alpha_1 \beta_a + \alpha_2 \gamma_a \tag{1}$$

The costs for the corresponding link of each travel mode are as follows:

(1) Car link

$$t_a = t_0^1 \left[1 + \alpha \left(\frac{v_a^2}{B_1 c_a} \right)^\beta \right] \quad (2)$$

$$\beta_a = F_1 d_a \quad (3)$$

$$\gamma_a = \rho_1 t_a \quad (4)$$

(2) Subway link

$$t_a = t_{w1}^2 + t_{w2}^2 \quad (5)$$

$$\beta_a = 0 \quad (6)$$

$$\gamma_a = \rho_2 t_a \quad (7)$$

(3) Bus link

$$t_a = t_{w1}^3 + t_{w2}^3 \quad (8)$$

$$\beta_a = 0 \quad (9)$$

$$\gamma_a = \rho_3 t_a \quad (10)$$

(4) Walking link

The walking link includes the transfer link and link entering the network.

$$t_a = t_{w1}^4 \quad (11)$$

$$\beta_a = 0 \quad (12)$$

$$\gamma_a = \rho_4 t_a \quad (13)$$

Related parameter definitions: t_a is the time cost of link a ; β_a is the economic cost; γ_a is the cost of comfort; α_1 and α_2 are the corresponding conversion coefficients; B_1 is the average number of people carried by the car, B_2 and B_3 are the maximum passenger capacity stipulated by the subway and bus respectively; $t_{w1}^2, t_{w1}^3, t_{w1}^4$ are the travel time of the corresponding travel mode on the link; t_{w2}^2 and t_{w2}^3 are the waiting

time of public transportation at the station; F_1 is the cost per unit mileage of the car; ρ_i is the comfort-time cost parameter corresponding to travel mode.

3 Model of Reserve Capacity

The definition of transportation network capacity in this paper is: under certain conditions of road and traffic control and level of service, for a given demand distribution structure, the urban traffic network can serve the largest number of travelers in a unit time.

3.1 Bi-Level Model

The upper-level model is to solve the maximum service volume under a given O-D pattern, where u is the multiplier of the O-D pattern, v_a is the link flow and c_a is the link capacity, Formula (15) shows the link capacity constraints for real links.

$$\max u \tag{14}$$

s.t.

$$v_a \leq c_a \tag{15}$$

The lower-level models is the description of traveler’s behavior, in this paper, we use logit-based SUE to describe traveler’s behavior, the random term of discrete route choice satisfies Gumbel distribution, the route choice probability is defined by:

$$P_k^{rs} = \frac{e^{-\theta c_k^{rs}}}{\sum_l e^{-\theta c_l^{rs}}} \tag{16}$$

When $\theta \rightarrow \infty$, it indicates that the traveler has a higher degree of grasp of road conditions and converges to UE assignment, the minimization program is the lower-level model:

$$\min \sum_a \int_0^{v_a} c_w(v)dv + \frac{1}{\theta} \sum_{r,s} \sum_k f_k^{rs} \ln(f_k^{rs}) \tag{17}$$

s.t.

$$f_k^{rs} \geq 0 \tag{18}$$

$$v_a = \sum_{r,s} \sum_k f_k^{rs} \delta_{a,k}^{rs} \tag{19}$$

In order to be more closer to the actual situation, taking into account factors such as ro-adside and intersection, the reduction formula for road capacity can be expressed as:

$$c_a = c \times f_1 \cdot f_2 \cdot f_3 \tag{20}$$

where n is the number of lanes, f_1 is the percentage of green light time at the entrance; f_2 is the roadside interference coefficient; f_3 is the road grade correction coefficient; C is the theoretical capacity.

3.2 Solution Algorithm

To solve this bi-level programming model, most studies use the descent method or design the related heuristic algorithms, when a bottle-neck link exceeds the capacity constraint, stop iterating. only a few link of the results are fully utilized. In order to make the network structure more fully utilized, a heuristic method based on incremental assignment technique is proposed to solve this model.

The upper-level model of the bi-level programming model has only one variable u , which is directly introduced into the lower-level model as a parameter. By increasing the value of u , repeatedly solve the lower-level models and return to the upper-level model constraints until the viable path set is empty. The calculation steps are given as following:

Step1: The traversal graph method determ

ines the viable path set K_{ij} between the OD pair (i, j) , which meets the level of service and restricts the number of transfers k between travel modes, $k \leq 1$:

$$t_k^{rs} \leq t_{rs}^{max}(\omega) = (1 + \omega)t_{rs}^{min} \tag{19}$$

Step2: Set the initial value of the multiplier of the demand distribution matrix u , $u^n = 1, n = 0$, and the incremental Δu .

Step3: Use the Method of Successive Weighted Averages (MSWA) [14] algorithm to assign the increment of u^n to get the link flow. The assignment steps are as following:

- ① On the basis of the existing assignment, take the zero-flow time and perform an alloc-ation to obtain v_a^m .
- ② Update the travel cost c_a^m of each link.
- ③ According to the updated travel cost for secondary assignment, the search direc-tion is determined, the additional traffic volume y_a^m of each link is obtained, and the search dire-ction $\Delta c_a^m = y_a^m - v_a^m$.

- ④ Update link flow. set $d = 1$.

$$v_a^{m+1} = v_a^m + \frac{m^d \Delta c_a^m}{1^d + 2^d + 3^d + \dots + m^d} \tag{21}$$

- ⑤ Convergence test. If the difference Δc_a^m meets the convergence requirement.

$$\frac{\sqrt{\sum (v_a^{m+1} - v_a^m)^2}}{\sum v_a^m} \leq \varepsilon \tag{22}$$

Then stop the calculation, otherwise, set $m = m + 1$ and return to Step ②.

Step4: If the current flow does not meet the restrictions (15), set $u^{n+1} = u^n + \Delta u$, and return to Step 3. Otherwise, cut the link that exceeds the constraint, and return to Step3 until there is an empty set of viable path sets K_{ij} , the solution is u^n .

4 Numerical Example

The construction of the test network is shown in Fig. 3, which includes two sets of OD pairs; three modes of transportation (car, subway and bus); two subway lines ($N_9-N_{10}-N_{11}$, $N_{12}-N_{10}$); two bus lines ($N_{13}-N_{14}-N_{16}-N_{19}$, $N_{17}-N_{18}-N_{16}-N_{15}$). Convert this network into a super network, the structure is shown in Fig. 2.

Assuming that the departure interval of the two subway schedules is 4 min/vehicle, the travel time of each link is 4 min, and the load limit is forty people, some parameters are given in Table 1. the given OD demand matrix is shown in Table 2, the travel time for each link is shown in Table 3 (Car) and Table 4 (Bus), the walking time for transfers between each mode is 2 min, and the waiting time of public transportation

Fig. 3 Test network

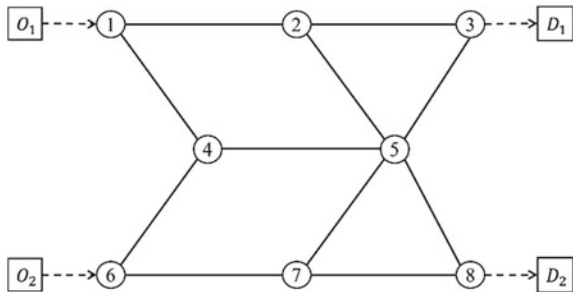


Table 1 Calculation parameters

α_1	α_2	F_1	F_2	ρ_1	ρ_2	ρ_3	ρ_4	θ	ω
0.05	0.1	4	10	0.2	0.3	0.6	0.7	1	0.5

Table 2 Initial OD Matrix

Origin/destination	D_1	D_2
O_1	300	200
O_2	200	300

Table 3 Parameters of car link

Link	7	8	9	10	11	12
t_a^0	7	5	4	6	4	5
C_a	250	260	200	210	360	300
Link	13	14	15	16	17	
t_a^0	4	5	5	7	5	
C_a	210	200	300	200	210	

Table 4 Parameters of bus link

Link	21	22	23	24	25	26
time	7	4	6	5	5	7

at the station is 1 min. It takes equal time to get in and out of the network. The bus schedules is 8 min/vehicle, and the load limit is forty people, Cars carry two passengers on average.

Search for the set of viable paths from the network: $K_{11}, K_{21}, K_{12}, K_{22}$, where the number of sets of each element is shown in Table 5.

Set the multiplier and increment initial value as follows: $u^0 = 1, \Delta u = 0.1$. The flow and capacity are counted as person, and the initial assignment results are shown in Table 6.

In the 9th assignment, the flow on A_{15} has exceeded the capacity, then cut the link, at this time, $u^8 = 1.8$, twelve paths are removed from the viable path set. The removal paths are shown in Table 7.

After 12th assignment, $n = 11, K_{12}$ become a null set, stop calculation, the final assignment result is shown in Table 8.

Finally, $u^{11} = 2.1$, that is, under the set demand structure, the transportation network can serve 2100 people within one hour. From the assignment process, the bottleneck link that limit the network capacity can be explicit found, in order: $A_{15}, A_{21}, A_{23}, A_{11}, A_{16}$, by expanding these links, the network capacity will get more promoted.

Table 5 Number of viable path

K_{11}	K_{21}	K_{12}	K_{22}
9	11	9	8

Table 6 Initial Results of Assignment

Link	v_a	v/c	Link	v_a	v/c
7	118.25	0.24	17	67.20	0.17
8	57.45	0.11	18	153.44	0.26
9	61.17	0.15	19	183.99	0.31
10	95.86	0.24	20	185.33	0.31
11	297.42	0.42	21	167.15	0.56
12	272.03	0.45	22	132.09	0.44
13	74.93	0.19	23	145.13	0.48
14	123.20	0.31	24	49.33	0.16
15	339.61	0.57	25	93.19	0.31
16	128.00	0.32			

Table 7 Disconnection Path

$N_1-N_2-N_5-N_8$	$N_6-N_7-N_5-N_8$
$N_1-N_4-N_5-N_8$	$N_6-N_4-N_5-N_8$
$N_{13}-N_{14}-N_2-N_5-N_8$	$N_6-N_7-N_5-N_8$
$N_9-N_{10}-N_{11}-N_5-N_8$	$N_6-N_{10}-N_4-N_5-N_8$
$N_9-N_{10}-N_4-N_5-N_8$	$N_6-N_{10}-N_{11}-N_5-N_8$
$N_{13}-N_{14}-N_{16}-N_5-N_8$	$N_{12}-N_{10}-N_4-N_5-N_8$

Table 8 Final Results Of Assignment

Link	v_a	v/c	Link	v_a	v/c
7	309.01	0.62	17	390.96	0.98
8	129.73	0.26	18	374.93	0.62
9	154.52	0.39	19	415.89	0.69
10	227.92	0.57	20	366.12	0.61
11	744.38	1.03	21	304.87	1.01
12	583.83	0.97	22	252.23	0.84
13	153.26	0.38	23	305.22	1.02
14	248.79	0.62	24	110.88	0.37
15	611.29	1.02	25	227.75	0.76
16	413.04	1.03			

5 Conclusions

This study set out to the bearing capacity of urban multimode transportation networks. Based on the combined travel mode, it gives the travel cost function, and constructs a bi-level programming model for solving the capacity of the multimode transportation

network and modifies the capacity. Taking into account the limitation of factors such as the number of transfers and level of service, the viable path set is restricted in the algorithm, this makes the result more accurate. At the same time, the network bottleneck link can be found in the process, and the expansion of the bottleneck link can obtain greater benefits within limited budget. However, it is necessary to deepen the research on the traveler's behavior to avoid 'illogical' routes, future studies should aim at the selection of viable path in the urban multimode transportation network.

Acknowledgements This study is supported by The National Key Research and Development Program of China (No. 2018YFB1600900).

References

1. Smeed RL (1996) Road capacity of city centers. *Traffic Eng Control* 8:455–458
2. Chen H, Zhou JB, Wu Y et al (2011). Modeling of road network capacity research in urban central area. *Appl Mech Mater* 1069:778–784
3. Sheffi Y (1985) *Urban Transportation Networks: Equilibrium analysis with mathematical program-ming methods*. Prentice Hall, Englewood Cliffs
4. Li P, Li B, Xie B et al (2009) Improved algorithm about the maximum flow in road network using auxiliary graph theory. *J Transp Syst Eng Inf Technol* 9(1):99–103
5. Yang, Michael GH, Bell, Meng M (2000) Modeling the capacity and level of service of urban transportation networks. *Transp Res Part B* 34(4):255–275
6. Chiou SW (2009) Optimization of limited network capacity with toll settings. *Inf Sci* 197:109–119
7. Baskan O (2019) A multiobjective bilevel programming model for environmentally friendly traffic signal timings. *Adv Civil Eng* 1–13
8. Suh-Wen C (2014) Optimal signal setting for road network with maximum capacity. *Inf Sci* 273:287–303
9. Fei H, Lin C (2016) Stochastic user equilibrium model with a tradable credit scheme and application in maximizing network reserve capacity. *Taylor & Francis* 49(4):549–564
10. Elnaz M, Reza ZF (2011) Optimizing reserve capacity of urban road networks in a discrete network design problem. *Adv Eng Softw* 42(12):1041–1050
11. Zhu J, Zhang N (2008) Modeling road network capacity and service level under variable demand pattern. *Syst Eng Theory Pract* 28(6):170–176
12. Wang, Deng, Zhao (2015) Road network reserve capacity with stochastic user equilibrium. *Transport* 30(1):103–116
13. Luo Z, Liu Y, Yu J (2012) Estimation of urban transportation network capacity considering traveler road preferences. *J Urban Planning Develop* 138(2):133–142
14. Liu HX, He X, He B (2009) Method of successive weighted averages (MSWA) and self-regulated averaging schemes for solving stochastic user equilibrium problem. *Netw Spat Econ* 9(4):485–503

Autonomous Train Dispatch Method for Regional Rail Transit Based on Deep Reinforcement Learning



Wei Wei, Ling Liu, Jun Liu, and Bo Zhang

Abstract Train dispatching of rail transit is an important issue for implementing the transportation plan and ensuring the orderly operation. Recognizing that existing centralized dispatch methods lack efficiency and flexibility, a novel concept of autonomous train dispatch of rail transit is proposed. Additionally, the Deep Reinforcement Learning (DRL) is incorporated into the autonomous train dispatch (ATD) model, and the resultant DRL-ATD model considers both dispatch and control objectives. The Double DQN (DDQN) algorithm is used for the model training, and a prioritized experience replay mechanism is further introduced to improve the stability and convergence. Using a case study involving the Chongqing Rail Transit in China, it is found that the DRL-ATD model can ensure the safety and punctuality of rail transit trains; it also reduces energy consumption by 1.62% during low passenger volume periods and passenger waiting time by 6.83% during high passenger volume periods, compared with the existing centralized method. Sensitivity analysis of the dispatch policy shows that the DRL-ATD model can effectively balance energy consumption of train operations and the total waiting time of passengers by adjusting the train operation schedules automatically, thereby providing the advantages of flexibility and dynamic adaptability for train operation.

Keywords Autonomous train dispatch · Deep reinforcement learning · Regional rail transit

1 Introduction

There are many advantages in rail transit system such as high speed, punctuality, saving land, etc. [1]. In such systems, train dispatching is an important issue for implementing the transportation plan and ensuring orderly operation. In existing transportation organization models, the transportation plan is usually pre-programmed

W. Wei (✉) · L. Liu · J. Liu · B. Zhang
Beijing National Railway Research & Design Institute of Signal & Communication Group Co.,
Ltd, Beijing 100073, China
e-mail: w_wenyun@qq.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_47

607

based on the passenger flow volume forecasted days or weeks in advance. However, fluctuations in real-time passenger flow result in a certain mismatch between transport supply and demand, and it leads to reductions in the service level and efficiency of the rail transit system. Furthermore, with the integrated planning and organization of multiple rail transit modes in a research region, such as urban agglomeration and metropolitan area, becoming a new trend, the existing dispatching methods are difficult to adapt to different modes of regional rail transit and the complex environments. It is therefore important to improve the efficiency and flexibility of rail transit with the help of intelligent train dispatch.

Optimization theory and methods are commonly used for train dispatch. Corman et al. [2, 3] proposed the use of the branch and bound algorithm and heuristic coordination algorithm to find the optimal solution for cooperative rescheduling and dispatch problems. Luis et al. [4] established a disturbance management model for rapid transit networks and applied GAMS and CPLEX to solve it. Törnquist and Persson [5] discussed train delay propagation resulting from disturbances in the railway network and discussed corresponding dispatch countermeasures. Cacchiani et al. [6] established a freight train scheduling and dispatching model using graph theory, and this improved the punctuality rate of rail transit trains. Dorfman and Medanic [7] proposed a local feedback-based travel advance strategy to handle perturbations in the schedule quickly. Albrecht and Panton et al. [8] proposed to use the Problem Space Search meta-heuristic model for large scale problems to recreate quality timetables.

From the above analysis, it can be seen that the existing research on train operations and dispatching focuses primarily on centralized models, i.e., they generate the dispatch scheme (adjusting the plan) for all operating trains simultaneously. However, during the course of train operations, the surrounding environment and transportation demands are extremely complex and exhibit strongly dynamic characteristics. In this case, the centralized dispatch methods are often lacking in efficiency and flexibility. Additionally, the dispatch problems of rail transit systems usually exhibit strongly nonlinear characteristics and real-time variability. The existing methods are not only restricted by low flexibility, but they also require long times to generate dispatch schemes.

In recent years, successful applications of reinforcement learning (DRL) in the dynamic decision-making field [9] have stimulated studies in the traffic and transportation fields. In the area of traffic control, Khamis and Gomaa [10] studied the use of multi-objective reinforcement learning for traffic signal control frameworks. Salkham et al. [11] proposed a phase switching model for collaborative reinforcement learning, which reduced intersection waiting time by approximately 57%. Jadhao et al. [12] combined the Q-learning algorithm with dynamic planning to make traffic control schemes adaptive to dynamic changes of traffic flow. In the field of autonomous vehicle driving, Sallab et al. [13] proposed a Deep Reinforcement Learning (DRL) framework for autonomous driving; this divided autonomous driving into the three stages: identification, prediction, and planning. Chae et al. [14] proposed a brake control system based on a Deep Q-learning Network (DQN) algorithm, which reduced the probability of vehicles hitting obstacles or pedestrians. Sallab et al. [15] proposed an end-to-end DRL method for lane assisted maintenance.

These models and methods provide important background for this paper. However, most of the existing studies focus on road traffic and rarely consider the characteristics and objectives of rail transit train dispatch. It is therefore difficult to apply them directly to rail transit train dispatch problems.

With the breakthroughs of the existing centralized mode, we propose herein a novel concept of autonomous train dispatch (ATD). Subsequently, a DRL based autonomous train dispatch model (DRL-ATD) is proposed, and both dispatch and control objectives are considered to improve the effectiveness and flexibility of rail transit transportation. Within DRL-ATD, the Double DQN (DDQN) algorithm is used for the model training to improve the stability and convergence, thereby solving the efficiency and flexibility problems that exist with existing dispatch methods. Finally, the effectiveness and feasibility of the proposed DRL-ATD model are verified with a case study involving the Chongqing Rail Transit in China.

2 Preliminary

In Reinforcement Learning, an agent is the decision-making subject that learns via “trial and error” to improve its own behaviour; it interacts with the environment and, in this way, finishes a given task [16]. During the trial and error process, an agent first interacts with the environment through tentative actions. As a result of the action, the environment will transform into a new state and give an immediate or delayed reward. The agent can improve its behavior policy with the reward from the environment and take action deemed appropriate for the new environment state. As the cycle continues, the agent interacts with the environment to gain experience and develop improved policies. After numerous iterations, the agent could learn the optimal action for completing the given task (i.e., the optimal policy to generate the optimal action) with a flexible response to the different state.

Using the ideas of Reinforcement Learning and the autonomous agent, we propose a novel concept of autonomous train dispatch (ATD). Figure 1 shows the principles of ATD. Each train adjusts its own dispatch action (i.e., the controlled acceleration) according to its own status and environmental conditions. We consider both dispatch objectives (punctuality and passenger waiting time) and control objectives (safety

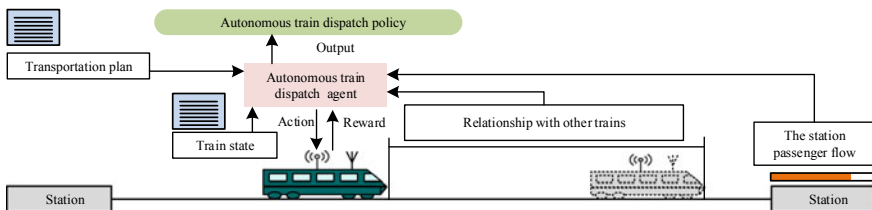


Fig. 1 Schematic diagram of ATD preliminary

and energy consumption) to improve effectiveness and flexibility of train dispatch. Compared to the traditional centralized mode, the ATD has the following advantages:

- The dispatch action is made for a single train and is more prompt and flexible than actions taken in centralized dispatch scenarios. The ATD mode can adapt to train dispatch tasks in complex scenarios, simplify dispatch system and reduce its construction costs.
- With ATD, the dispatch scheme is self-generated and implemented immediately rather than computed by the dispatching center and then transferred to the operating train. In this way, the likelihood of time delay and information loss during transmission is reduced, thereby improving the reliability of train dispatching.

For the convenience of the autonomous train dispatch modeling based on Reinforcement Learning, the following assumptions are introduced:

- The flow of waiting passengers at the station platform can be predicted accurately before the train arrives at the station.
- Information on the state of the rail transit system can be accurately obtained during the train's operation, regardless of errors in the collection and transmission of system information.

3 DRL-ATD Model

DRL is comprised of Reinforcement Learning and Deep Learning. It is widely used to solve the dynamic decision-making problem with continuous state space, for which conventional Reinforcement Learning is not applicable. We propose a DRL-based autonomous train dispatch model (DRL-ATD) in this paper. In the process of the action-value function fitting, the widely used DQN generally suffers from the problem of overestimation [17]. In other words, the action-value function obtained with the DQN approximation is larger than is the expected result. Since the Double DQN (DDQN) [18] method can effectively solve the overestimation problem of the conventional DQN, we choose the DDQN algorithm to train the proposed DRL-ATD model.

3.1 Reward Function

Taking the single train as research object, the state of the rail transit system and the dispatch action must first be determined according to the structure of DRL. The system state S_t at time t is the basis for the decision of dispatch action:

$$S_t = [L_t, P_{S_t}, P_{T_t}, T_t, l_t, y_t, v_t, z_t, \sigma_t] \quad (1)$$

where L_t denotes the operation time needed to go from the current position of the train (with the current speed) to the position of another train in front of it, P_{S_t} represents the predicted waiting passenger volumes at the forthcoming stations at time t and $P_{S_t} = [ps_t^1, \dots, ps_t^h, \dots, ps_t^H]$, ps_t^h represents the waiting passenger volume at the forthcoming station h ($h \in \{0, 1, 2, 3, \dots, H\}$), P_{T_t} is the residual capacity of the subject train at time t , T_t represents the total operating time of the subject train since departure, l_t represents the rail transit line of the research train at the time t (one-hot coded), y_t denotes the position of the subject train on the rail transit line, v_t and z_t are the running speed and acceleration of the studied train at the time t , respectively, and σ_t indicates whether the train is stopping at a station at the time t (0 stands for no and 1 for yes).

We take the controlled acceleration of the train as the autonomous dispatch action A_t . With the appropriate adjustment of dispatch action, the safety and punctuality of the operating train can be guaranteed, and the energy consumption and passenger waiting time are reduced.

As a result of autonomous dispatch action A_t with state S_t , the environment can give a reward to the train as a guide for improving autonomous dispatch policy. Taking both dispatch and control objectives into account, the reward function $R(S_t, A_t)$ is designed as:

$$R(S_t, A_t) = -[D(S_t) + C(A_t) + F(S_t) + B(S_t)] \tag{2}$$

where $D(S_t)$ is the waiting time cost function for passengers, $C(A_t)$ represents the energy consumption cost function for train operation, and $F(S_t)$ and $B(S_t)$ denote the safety cost function and the punctuality cost function for train operation.

(1) Waiting time cost function

The waiting time cost function $D(S_t)$ can be calculated based on the predicted passenger volume at the forthcoming stations of the subject train, as in Eq. (3):

$$D(S_t) = \omega \cdot L_t \cdot \sum_h^H \alpha^h ps_t^h / \sum_h^H \alpha^h \tag{3}$$

where ω is the economic cost per unit of passenger waiting time and α represents the discount rate for waiting time cost, $\alpha \in (0, 1)$.

(2) Energy consumption cost function

For the widely used rail transit trains, there are n discrete levels of controlled acceleration. The controlled acceleration level of the train at the time t is denoted by integer values of ψ_t , a positive value representing acceleration and a negative value indicating deceleration. More specifically, we suppose that A_t represents the transition from the controlled acceleration level ψ_t at time t to the next level ψ_{t+1} at time $t+1$. Therefore, the energy consumption cost function can be formulated as:

$$C(A_t) = \begin{cases} \delta(A_t) \cdot K + [1 - \delta(A_t)] \cdot (M \cdot z_t \cdot v_t \cdot \Delta t \cdot \varepsilon + \partial(A_t) \cdot b_{ij}) & \text{if } z_t > 0 \\ \delta(A_t) \cdot K + [1 - \delta(A_t)] \cdot \partial(A_t) \cdot b_{ij} & \text{else} \end{cases} \quad (4)$$

where M denotes the train mass and $z_t \cdot M \cdot v_t$ is the train traction power at time t , Δt is the time length of each discrete step for decision-making, ε is the economic cost per unit of energy consumption, K represents a large positive value (like 10^4), and b_{ij} is the economic cost per transition time of the controlled acceleration level. The function $\delta(A_t)$ indicates whether the transition from controlled acceleration level ψ_t to ψ_{t+1} is allowed based on the train's operating stability and passenger comfort. If this transition is not allowed, $\delta(\psi_t, \psi_{t+1})$ takes the value of 1; otherwise $\delta(\psi_t, \psi_{t+1})$ is 0. $\partial(A_t)$ represents whether the controlled acceleration level ψ_t and ψ_{t+1} are the same.

If the controlled acceleration level does not change from time t to $t + 1$, the energy consumption cost for train operation is $z_t \cdot M \cdot v_t \cdot \Delta t \cdot \varepsilon$ for acceleration motion and 0 for constant speed and deceleration motion. If the controlled acceleration level changes during the time increment t to $t + 1$, the energy consumption cost for the operating train is the sum of the traction energy and the level transition for acceleration motion ($\delta(A_t) \cdot K + [1 - \delta(A_t)] \cdot (z_t \cdot M \cdot v_t \cdot \Delta t \cdot \varepsilon + b_{ij})$) or the acceleration level transition cost only for even speed and deceleration motion ($\delta(\psi_t, \psi_{t+1}) \cdot K + [1 - \delta(\psi_t, \psi_{t+1})] \cdot b_{ij}$).

(3) Safety cost function

The safety cost function $F(S_t)$ is formulated as:

$$F(S_t) = \begin{cases} K \cdot \chi & \text{if } L_t < Md \\ \frac{\chi}{L_t - Md} & \text{else} \end{cases} \quad (5)$$

where L_t and Md represent the actual and minimum time distance between the subject train and the train in front of it, respectively, at time t , χ is the weight of a penalty for unsafe operation and K is a large positive number representing the penalty assessed for unsafe operation.

If the time distance at time t between the subject train and the train in front of it is less than the minimum value Md , the safety cost is set as a large value to assess a penalty for unsafe operation. Otherwise, the safety cost decreases with the increase of excess time distance ($L_t - Md$).

(4) Punctuality cost function

The punctuality cost function $B(S_t)$ for the autonomously dispatched train at the system state S_t is as:

$$B(S_t) = \begin{cases} K \cdot \xi & \text{if } T_t < Tmi(l_t, y_t) \\ K \cdot \xi & \text{if } T_t > Tma(l_t, y_t) \\ 0 & \text{else} \end{cases} \quad (6)$$

where T_t represents the total running time of the train from departure, $Tmi(l_t, y_t)$ and $Tma(l_t, y_t)$ are the shortest and the longest total running times allowed for the subject train to arrive at the next station, respectively, and ξ denotes the weight of the penalty for unpunctuality.

If a train arrives at the station between the earliest time $Tmi(l_t, y_t)$ and the latest time $Tma(l_t, y_t)$, the punctuality cost for it is 0. Otherwise, the punctuality cost is set as a large value for the unpunctuality penalty. In this way, the arrival times of the train at all stations can be constrained within an acceptable range.

3.2 Model Training based on DDQN

(1) Action exploration mechanism

The ε -greedy strategy is used to explore autonomous dispatch actions and collect the experience to improve the autonomous dispatch policy. Under the effects of an exploratory dispatch action A_t at system state S_t , the environment can give a reward $R(S_t, A_t)$ to the train, and the system state then transforms to state S_{t+1} . Removing the time information, we get the transition sample $i (S_i, A_i, R_i, S'_i)$ composed of the current system state S , the current action A , the current reward R , and the next system state S' . All the explored transition samples constitute the experience data set.

In ε -greedy strategy, a number within the interval $(0, 1)$ is first randomly generated. If this random number is less than ε , an autonomous dispatch action is randomly selected from all possible actions under the rail transit system state and executed to obtain the reward function and cause transition to the next state. Otherwise, the optimal action \bar{A} with the maximum value function is selected as the current action A .

(2) Priority replay mechanism

In DDQN, different models are used for action selection and action evaluation. The current action-value function Q is adopted for action selection, while the target action-value function Q^- is adopted for action evaluation. The action-value functions Q and Q^- are fitted by a neural network and the time-difference target Y_t^{DDQN} is defined as:

$$Y_t^{DDQN} = R(S_t, A_t) + \gamma Q^- \left(S_{t+1}, \arg \max_{a'} Q(S_{t+1}, a'; \theta); \theta^- \right) \quad (7)$$

where γ is the discount coefficient of the reward function and is a positive real number within the interval $(0, 1)$, θ and θ^- are learnable parameters.

When training the deep learning neural network for Q and Q^- , it is usually assumed that the training data are independent and distributed. However, there is a strong correlation in the data collected through ε -greedy strategy, and the neural

network is very unstable if these data are used for training. Therefore, it is necessary to use the method of experience replay to sample the transition (S_i, A_i, R_i, S'_i) from the experience data set.

Prioritized Experience Replay is a widely used sampling method [19, 20]. It can effectively improve the utilization of the explored data by assigning higher selection probabilities to the samples with more learning information. The sampling weight used for Prioritized Experience Replay is as:

$$p_i = |\delta_i| + \mu \quad (8)$$

where δ_i is the evaluation error, $\delta_i = Y_i^{DDQN} - Q(S_i, A_i)$, μ is a minimum positive number needed to avoid $p_i = 0$.

A higher absolute value of δ_i indicates that there is more learning information in the transition sample i , and the sampling probability of the transition sample i can be calculated as:

$$P(i) = \frac{p_i^\alpha}{\sum_n p_n^\alpha} \quad (9)$$

where α is the prioritized replay factor, a value of 0 means no prioritized replay, and 1 means completely prioritized replay, N is the size of the current experience data set.

Generally, the probability distributions of the experience data and the action-value function are different. To compensate for this bias, we introduce the importance of sampling weight w_i as:

$$w_i = \left(\frac{1}{N} \cdot \frac{1}{P(i)} \right)^\beta \quad (10)$$

where β is the compensation coefficient for importance sampling, a value of 0 means that the importance of sampling compensation has no effect, and 1 means that the compensation is effective.

A five-layer fully connected neural network is used to describe the action-value function for the ATD problem studied here. This network includes the input layer, output layer, and 3 hidden layers. The input is the current state S , and the output is the set of action-values for all possible actions. As mentioned above, the training of the action-value function network in DDQN is very different from that in DQN. The detailed algorithm process for DDQN can be seen in reference [18]. Specifically, we assume that the parameter set of the deep natural network to fit the action-value function is θ . According to the neural network structure, the parameter θ can be updated with the transition sample i :

$$\theta \leftarrow \theta + (Y_i^{DDQN} - Q(S_i, A_i)) \cdot \nabla_\theta Q(S_i, A_i) \quad (11)$$

where $Q(S_i, A_i)$ is the current action-value and $\nabla_{\theta} Q(S_i, A_i)$ is the relevant gradient function.

Considering the Prioritized Experience Replay, we should modify this update formula as:

$$\theta \leftarrow \theta + \eta \cdot w_i \cdot \delta_i \cdot \nabla_{\theta} Q(S_i, A_i) \tag{12}$$

where η is the update rate of θ . It is designed to decrease with the training process to facilitating convergence.

4 Case Studies

To verify the feasibility and effectiveness of the proposed DRL-ATD model, a case study is conducted using the Chongqing Rail Transit Line 1 in China. Figure 2 shows the scope of the case study, which includes 5 stations (and 1 depot) and 5 sections. The time length of a decision step for the autonomous train dispatching agent is set as 2 s. Each study time t actually represents a time period of 2 s. The operation plan studied is a parallel train diagram with a 3-min departure interval and the minimum safe time distance between any two adjacent trains is 1.5 min. The allowable deviation in the punctuality of a train arriving at the station is 1 min. The economic cost coefficient for rail transit train energy consumption is ¥1.025 per kilowatt-hour and the economic cost for passenger waiting time is ¥0.1254 per minute. The value of K is set as 10^4 to assess penalties for unsafe and unpunctual operations. The weights χ and ξ for these penalties are all 1.0.

In order to realize the precise dispatch of the autonomous train with discrete actions, the controlled acceleration levels are designed with a proportional structure based on the widely used 15-levels acceleration for the urban rail transit trains of China. Table 1 shows the configuration of train acceleration levels used here.

For the study case, the proposed DRL-ATD model is trained with both the DDQN and DQN algorithms. Figure 3 shows the change curves for the optimal action-value function (Q value for the optimal dispatch action) with the DDQN and DQN training processes, respectively. Compared with DQN, the DDQN algorithm converges faster and also has a more stable action-value function. With different models used for action selection and action evaluation and with the introduction of Prioritized Experience Replay, DDQN improves the training efficiency of the DRL-ATD model. It

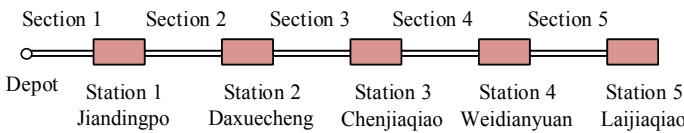


Fig. 2 Scope of the case study

Table 1 Acceleration levels of autonomous dispatch train

Acceleration level	Acceleration (m/s^2)	Acceleration level	Acceleration (m/s^2)	Acceleration level	Acceleration (m/s^2)	Acceleration level	Acceleration (m/s^2)	Acceleration level	Acceleration (m/s^2)
7	1.000	3	0.010	-1	-0.001	-5	-0.100		
6	0.300	2	0.003	-2	-0.003	-6	-0.300		
5	0.100	1	0.001	-3	-0.010	-7	-0.700		
4	0.030	0	0.000	-4	-0.030				

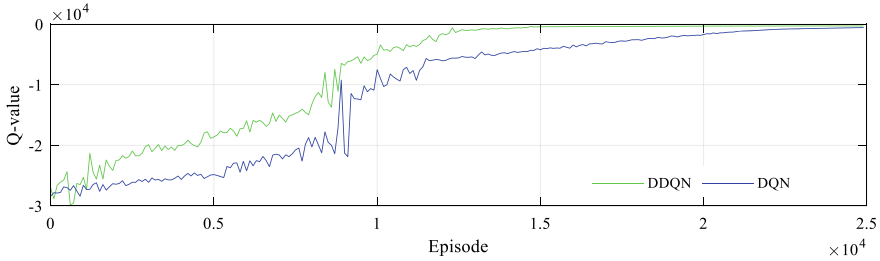


Fig. 3 Training process comparison for DQN and DDQN

is thus more applicable to the complex decision-making problem of ATD than is the conventional DQN.

Furthermore, it can be seen from Fig. 3 that the optimal Q values are far greater than -10^4 at the end of the training. This illustrates that there are no penalties assessed for unsafe or unpunctual operation with the trained autonomous dispatch policy. The proposed DRL-ATD model can ensure the safety and punctuality of rail transit trains.

Figure 4a shows the variation in train energy consumption when the passenger volume at Station 1 decreases by 50%. With the low passenger volume of Station 1, the model can automatically adjust the train speed at Sect. 1 to reduce the energy consumption of the train. In the subsequent sections, the running speed is improved to ensure train punctuality and maintain passenger waiting time within an acceptable range. With the proposed DRL-ATD model, the total energy consumption is reduced by 1.62% in contrast to the existing centralized dispatch method.

Figure 4b shows the change in passenger waiting time when the passenger volume of Station 1 increases by 50%. The DRL-ATD model can automatically raise the running speed of the train in Sect. 1 to reduce the passenger waiting time at Station 1. The total passenger waiting time with the DRL-ATD model is 6.83% less than is that with the existing centralized dispatch method.

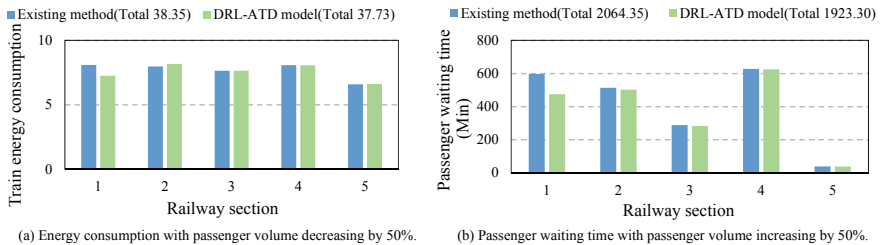


Fig. 4 Energy consumption and passenger waiting time under changing passenger flow

5 Conclusions and Future Work

In this paper, we proposed a novel concept of autonomous train dispatch (ATD) to solve the inefficiency and inflexibility problems existing in centralized dispatch methods. Considering both dispatch and control objectives, the DRL-ATD model is proposed based on Deep Reinforcement Learning. The DDQN algorithm is used for model training, and the Prioritized Experience Replay mechanism is further introduced to improve the stability and convergence of the model training. In the case study using data from the Chongqing Rail Transit Line 1 in China, the proposed DRL-ATD model with DDQN exhibits better convergence and stability than does the DQN algorithm. The proposed model can ensure the safety and punctuality of rail transit trains, and it also reduces energy consumption by 1.62% during periods of low passenger volume and reduces passenger waiting time by 6.83% during periods of high passenger volume. By considering the dynamic transport demands, the proposed DRL-ATD model can automatically balance energy consumption and passenger waiting time. This shows the superior flexibility and dynamic adaptability for train operation.

With Continuously Variable Transmission becoming an important tool for the operational control of modern rail transit trains, in the future work, we will further investigate ATD methods based on other Reinforcement Learning models with continuous action space.

References

1. Bu B, Guoying Q, Ling L, Guojie L (2018) An energy efficient train dispatch and control integrated method in uUrban rail transit. *Energies* 11(5):1248
2. Corman F, Ariano AD, Pacciarelli D, Pranzo M (2012) Optimal inter-area coordination of train rescheduling decisions. *Transp Res Part E Logist Transp Rev* 48(1):71–88
3. Corman F, D' Ariano A, Pacciarelli D, Pranzo M (2014) Dispatching and coordination in multi-area railway traffic management. *Comput Oper Res* 44:146–160
4. Luis C, Ángel M, Gábor M (2013) Recovery of disruptions in rapid transit networks *Transp Res Part E Logist Transp Rev* 53(1):15–33
5. Rnquist JT, Jan AP (2007) N-tracked railway traffic re-scheduling during disturbances. *Transp Res Part B* 41(3):362
6. Cacchiani V, Alberto C, Paolo T (2004) Scheduling extra freight trains on railway networks. *Transp Res Part B Methodol* 44(2):231
7. JDM, MJ (2004) Scheduling trains on a railway network using a discrete event model of railway traffic. *Transp Res Part B (Methodol)* 38(1):98
8. Albrecht AR, Panton DM, Lee DH (2013) Rescheduling rail networks with maintenance disruptions using Problem Space Search. *Comput Oper Res* 40(3):703–712
9. Mnih V et al (2015) Human-level control through deep reinforcement learning. *Nature* 518(7540):529–533
10. Khamis MA, Gomaa W (2014) Adaptive multi-objective reinforcement learning with hybrid exploration for traffic signal control based on cooperative multi-agent framework. *Eng Appl Artif Intell* 29:134–151

11. Cunningham R, Garg A, Cahill V (2008) A collaborative reinforcement learning approach to urban traffic control optimization. In: Proceedings. 2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology
12. Jadhao NS, Jadhao AS (2004) Traffic signal control using reinforcement learning. In: International conference on communication systems & network technologies
13. Sallab A, Abdou M, Perot E, Yogamani S (2017) Deep Reinforcement Learning framework for Autonomous Driving. *Electronic Imaging 2017*(19):70–76
14. Chae H, Kang CM, Kim B, Kim J, Chung CC Choi JW (2017) Autonomous braking system via deep reinforcement learning. In: Proceedings of IEEE international conference on intelligent transportation systems
15. Sallab AE, Abdou M, Perot E, Yogamani S (2016) End-to-end deep reinforcement learning for lane keeping assist. arXiv: 1612.04340
16. Sutton R, Barto A (1998) Reinforcement learning: an introduction. *IEEE Trans Neural Netw* 9(5):1054–1054
17. Anshel O, Baram N, Shimkin N (2016) Deep reinforcement learning with averaged target DQN. arXiv: 1611.01929v1
18. Van Hasselt H, Guez A, Silver D (2015) Deep reinforcement learning with double q-learning. *Computer Science*
19. Horgan D et al (2018) Distributed prioritized experience replay. arXiv: 1803.00933
20. Schulze C, Schulze M (2018) ViZDoom: DRQN with prioritized experience replay, double-q learning, & snapshot ensembling. arXiv: 1801.01000

Research on Energy Delivery of a Series–Parallel Hybrid Electric Vehicle Under Different Driving Conditions



Chunming Li, Xiaoxia Sun, Chunming Shao, Lining Yang, Chenglong Shu, Danhua Niu, and Guozhu Wang

Abstract In series–parallel hybrid electric vehicle, the relationship between power flow and heat flow is very complex. The dynamic thermal characteristic of series–parallel hybrid electric vehicle is closely related to vehicle energy management strategy, external environment and driving conditions. In this paper, the matching of the vehicle working mode with the vehicle operating conditions is studied. The theoretic analysis of vehicle energy distribution under different typical operating condition is conducted. A coupled whole vehicle simulation model which is built in MATLAB/SIMULINK and GT suite is proposed. According to the theoretical qualitative analysis and simulation quantitative analysis, the corresponding relationship combined heat response and power consumption between power flow and heat flow is acquired. Study results show that the coupled simulation model can be effectively used in the research combined dynamic heat response, the power and heat delivery and power consumption under different operating conditions and working modes. This integrated research method combined theoretical qualitative analysis and simulation quantitative analysis is an efficient way in the multi-disciplined complex vehicle system energy management study.

Keywords Energy distribution · Series–parallel hybrid electric vehicle · Power and heat delivery · Multi-system modelling

1 Introduction

Nowadays, concerns focused on climate change, air pollution and future of energy issues are dramatically increased around the world. The hybrid electric vehicle is a good solution to settle down the problems and a suitable alternative to conventional internal combustion engine vehicle. Series–parallel hybrid electric vehicle (SPHEV) has the advantages of both the series hybrid electric vehicle and the parallel hybrid electric vehicle in technical, economic and environmental aspects. As the GM dual

C. Li · X. Sun (✉) · C. Shao · L. Yang · C. Shu · D. Niu · G. Wang
China North Vehicle Research Institute, Beijing, China
e-mail: sun_xiaoxia1983@163.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_48

621

mode hybrid electric vehicle and Toyota Prius appearing on the world market, there are a lot of researches on the hybrid electric vehicle with the coupled planetary mechanism of split and convergence characteristics.

Series-parallel hybrid electric vehicle system includes mechanical, electrical, thermal management, control and other subsystems. There are multiple coupling relationships between each other. Under a complex driving condition, the vehicle powertrain interacts with the thermal management system, forming the power flow and heat flow multiple coupling influence relationship. Therefore, it is necessary to master the coupling relationships among the different parameters in different subsystems by the deep analysis of interactions between the power flow and heat flow. At last, the multi-parameter influencing analysis system integrated ambient environment, design schematic and energy management strategy can be formed. The multi-disciplinary coupling model of the series-parallel hybrid electric vehicles can be established.

There is a complex relationship between heat flow and power flow in series-parallel hybrid electric vehicle illustrated in Fig. 1. Previous studies often focus on one aspect while neglecting the other one especially in the series-parallel hybrid electric vehicle [1-4]. The dynamic thermal characteristic of series-parallel hybrid electric vehicle is closely related to vehicle energy management strategy, external

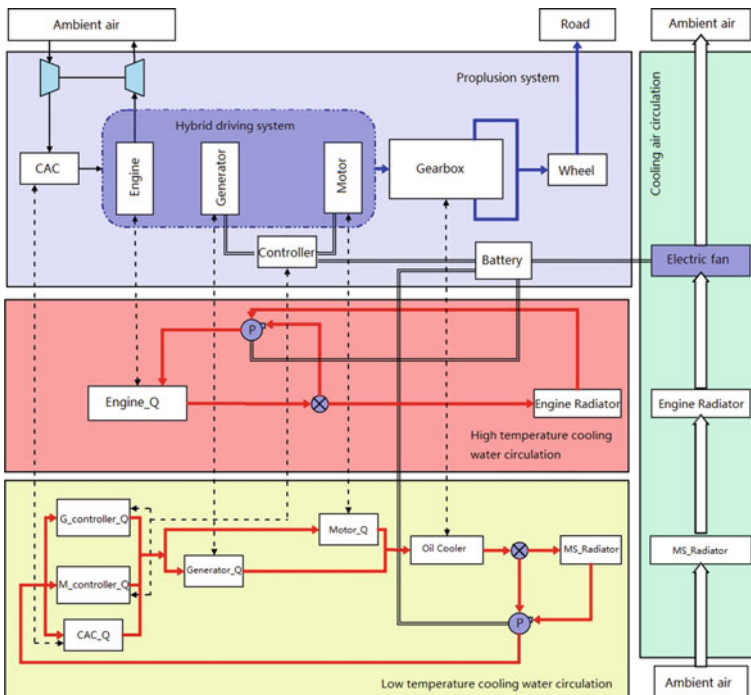


Fig. 1 Schematic of coupling relationship between power flow and heat flow in the propulsion system and thermal management systems

environment and driving conditions. In order to improve the economy performance of the vehicle and reduce the energy consumption, it is necessary to define the energy demand and conversion process in the vehicle system. Based on the energy management strategy, the multi-disciplinary integrated simulation is conducted to study the power and heat delivery under different operating conditions and working modes.

2 Schematic of Series-Parallel Hybrid Electric Vehicle

A. The schematic of hybrid driving system

The schematic of the series-parallel hybrid driving system is illustrated in Fig. 2. The hybrid power system of series-parallel hybrid electric vehicle includes the internal combustion engine, the motor A, the motor B, the planetary power coupling device and so on. The internal combustion engine is linked to the planet carrier, the motor A is linked to the sun gear, and the motor B is linked to the ring which is paralleled on the output shaft. Both the motor A and the motor B can operate as generators or driving motors. In the series-parallel hybrid electric vehicle, the engine power is split by the planetary power coupling device. Some of the engine power is transmitted to the output shaft as mechanical power. Some of the engine power is transmitted by the motor A to the motor B or battery as electrical power. The torques of the motor B and the planet ring are coupled on the output shaft, and power split and coupling are implemented by the planetary power coupling device.

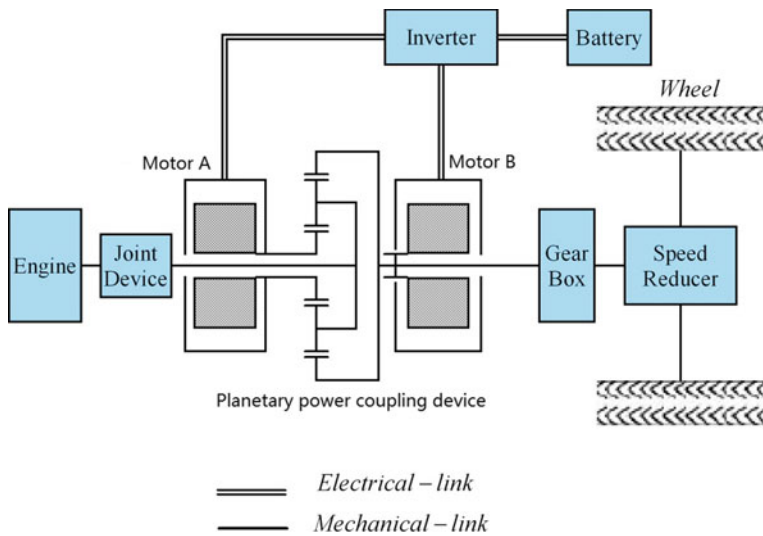


Fig. 2 Schematic of series-parallel hybrid electric vehicle with planetary power coupling device

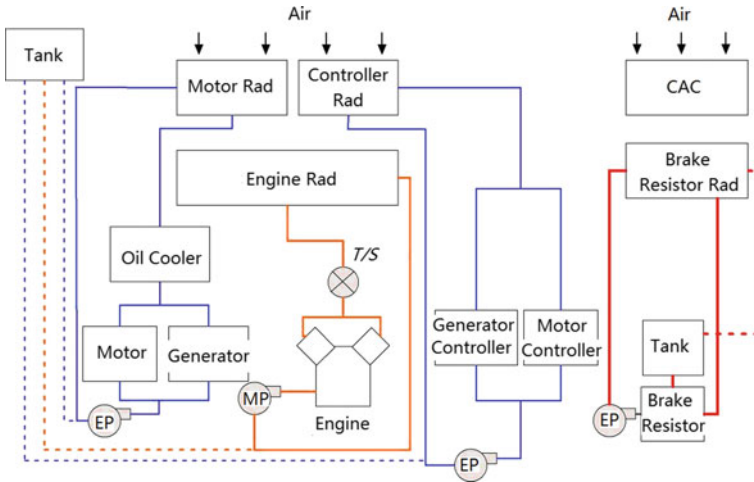


Fig. 3 Schematic of power-split hybrid electric vehicle thermal management system (Rad: Radiator, EP: Electric Pump, MP: Mechanical Pump, T/S: Thermostat, CAC: Charge Air Cooler)

B. The schematic of thermal management system

The thermal management system is designed with two separate air duct systems illustrated in Fig. 3. Two electrical fans draw air through two different cooling stack air pathways which can work synchronously or not depending on the different driving conditions. It has four fluid-cooling circulation loops. One cooling fan draws air through the cooling stack air pathway removing waste heat from controller cooling circuit, motor cooling circuit and engine cooling circuit in succession. The controller cooling circuit and the motor cooling circuit are parallel arranged. The other one cooling fan draws air through the cooling stack air pathway removing waste heat from charged air cooling circuit and brake resistor cooling circuit in succession. In electric power consuming braking process, the braking energy is turned into electrical energy to charge the batteries. The brake resistor sub-system is used to consume the remaining electricity once the electrical energy is still surplus after the recharging process.

3 Analysis of Power and Heat Delivery of Series-Parallel Hybrid Electric Vehicle

A. The working mode of hybrid electric vehicle

The hybrid driving system of series-parallel hybrid electric vehicle can realize a variety of working mode. The study based on working modes is the premise for the analysis of power and heat interacting delivery. In order to realize the optimized

performance of the hybrid driving system, the working modes need to match the different operating conditions.

Electric driving mode. This working mode is applied in the low speed and small load condition. In this situation, output shaft speed is lower than the minimum engine speed. In this working mode, the battery is fully charged. The driving motor can fulfill the power requirement of vehicle driving and thermal management system consumption. In this drive mode, the maximum available vehicle torque is limited by the maximum torque of the motor itself and the maximum power which battery can supply.

Engine single driving mode. This working mode is applied in the situation that vehicle required torque is in the engine high efficiency load areas. Depending on the battery state of charge (SOC), the battery pack is selected to operate in either the charge mode or the discharge mode. When the SOC is less than the target value, the battery pack needs to be charged. Under this circumstance, the generator plays roles that not only adjust the engine speed but also charge the battery pack. When the SOC value is very low, the urgent needing to be charged urgently, the motor also works as a generator.

Hybrid driving mode. This working mode is applied to full load acceleration or climbing conditions. When the vehicle demand torque is greater than the maximum torque of the engine, the generator plays roles to adjust engine speed as well as exporting electric power such as drive thermal management system or drive the driving motor.

Braking mode. Braking mode includes electrical braking and mechanical braking. Electrical braking can be divided into electric regenerative braking mode and electric energy consumption braking. If the battery SOC reaches its upper charge limit, the electric regenerative braking need to be changed into electric energy consuming brake. In this situation, active the brake resistor, transforming the electricity into heat. The surplus electricity can be fully used according to the heat circulation such as heating cabin and warming the engine and so on.

Vehicle stopping and charging mode. This working mode is a special case of the vehicle stopping mode. In this mode, the transmission gear is set at neutral position. Engine is used to drive the generator and motor through the planetary coupling mechanism power. Under this circumstance, the motor used as a generator to charge the battery pack, or power the external load.

B. *The relationship between the vehicle driving mode and driving condition*

In the operation of series-parallel hybrid electric vehicle, the matching of vehicle driving mode and driving condition is very important. The corresponding relationship and energy distribution with the matching of typical driving condition and basic working mode of PSHEV are shown in Fig. 4. The segments AB, IJ and JK are all vehicle stop conditions. The segments AB and IJ apply stop mode, while segment JK applies vehicle stop but charging mode. The segment BC is vehicle electric starting and speeding up condition which applies electric driving mode. The segment CD

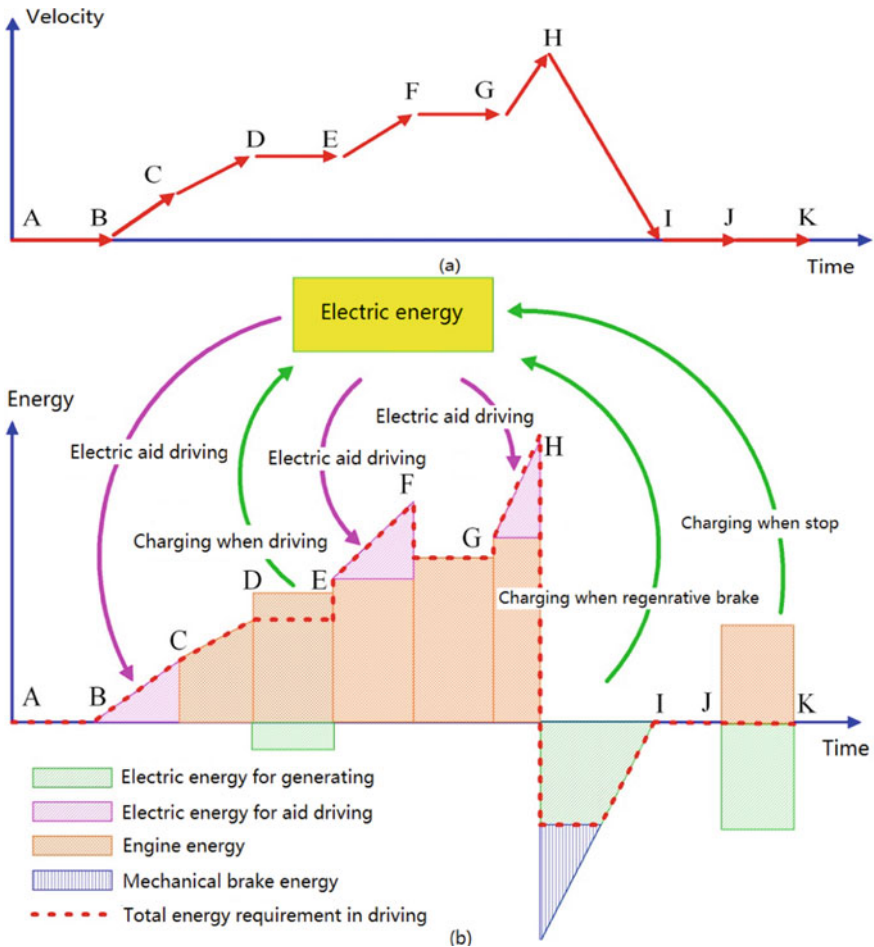


Fig. 4 The energy distribution with the matching of typical driving condition and basic working mode

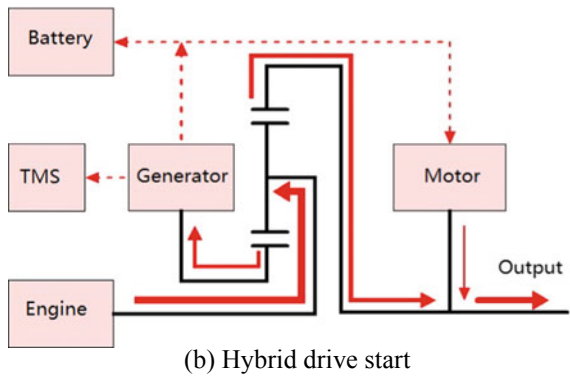
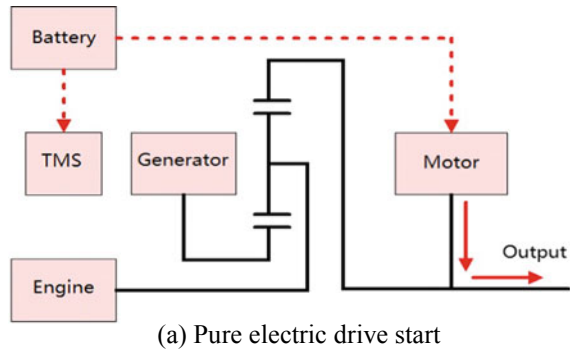
is single engine driving and acceleration condition, which uses the hybrid driving mode. Under this circumstance, the acceleration is relatively small. The segments DE and FG are vehicle high-speed cruising conditions, which use hybrid driving mode. The differences of the two segments are the different SOC states. The segment DE uses engine independent driving with charging mode, while the segment FG applies engine independent driving without charging mode. The segments EF and GH are vehicle accelerating conditions, which use hybrid driving mode. The segment HI is vehicle deceleration/brake condition which uses brake mode. In this situation, the driving motor works as a generator, supplying the brake force as well as charging. Under this condition, the distribution of brake force generated by the mechanical brake and electrical brake should follow the changing of the brake strength.

C. Analysis of energy distribution under typical driving conditions

Vehicle start condition. For series-parallel hybrid electric vehicle, there are two start forms- pure electric drive start and hybrid start. When starting on a good road, vehicle can be driven by electric motor only. After the velocity reaches to a certain value, the engine begins to work. The energy distribution under vehicle starting condition is shown in Fig. 5a. Pure electric drive starting can reduce the frequency that engine works in low load and low speed condition, so as to improve the overall operating efficiency of the engine. It is very important for reducing the vehicle’s fuel consumption and emissions.

When starting on a tough road or the battery SOC is low, the maximum motor torque cannot meet the power needs of the start. In this situation, the engine needs to start to achieve hybrid drive start. The energy distribution under vehicle starting condition is shown in Fig. 5b. In this operating condition, the output torque of the vehicle can increase steadily by adjusting the output torque of the generator. At this time, the engine drives the generator to generate and drive the vehicle at the same time in order to ensure the engine works in its efficient working area. If the conditions permit, the pure electric drive start is always used preferentially in most of time.

Fig. 5 The energy distribution under vehicle starting condition (TMS: Thermal Management System)



Vehicle cruising condition. The cruising condition of series–parallel hybrid electric vehicle includes low-speed cruising and high-speed cruising. Low-speed cruising condition uses the electric drive mode. The high-speed cruising condition employs the engine single driving mode. High-speed cruising is a driving condition in which the vehicle aims at maintaining a medium–high driving target speed by adjusting output power with the external road and so on. In this condition, motor is not used or works as generating condition. This condition can be divided into battery discharging–engine alone driving and battery charging–engine alone driving two kinds of situations. When the road loads change, there is no need to adjust the state of hybrid powertrain. Instead of that, it needs to adjust the engine throttle and generator output torque in order to achieve the changing of the vehicle output torque. In the vehicle operating process, the battery working state–charging or discharging, rely on the battery SOC state. When the SOC is lower than the target value, the generator and motor can be used to charge the battery pack. If the SOC is higher than the target value, the battery pack will discharge according to the vehicle power demand. Under this circumstance, it is possible to keep the engine operating in an efficient area so as to ensure the high efficiency and low emissions of the vehicle. The energy distribution under vehicle cruising condition is shown in Fig. 6.

Acceleration/climbing condition. These operating conditions are the vehicle’s extreme conditions. The demanding power of the vehicle driving is very huge. In the situation, vehicle employs the mechanical and electrical hybrid driving mode. The engine drives the vehicle combined with the battery. The specific energy distribution can be seen in Fig. 7.

Fig. 6 The energy distribution under vehicle cruising condition (TMS: Thermal Management System)

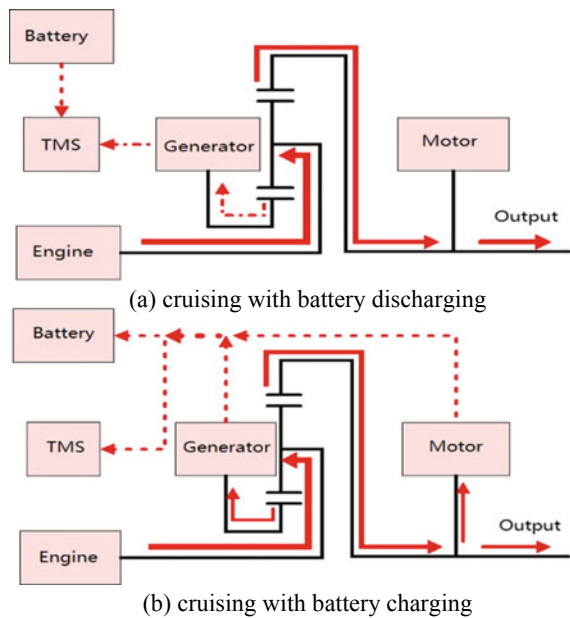
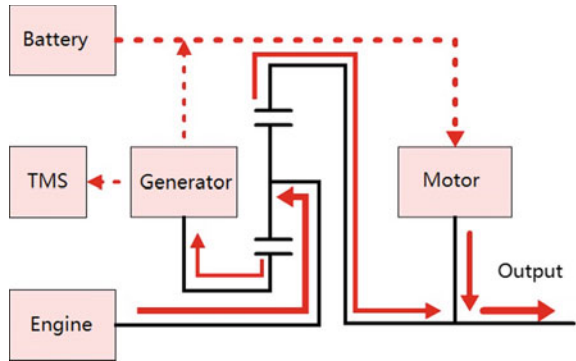
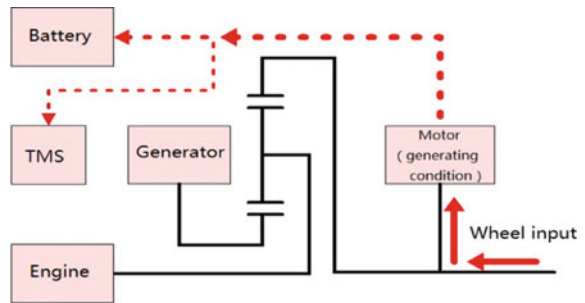


Fig. 7 The energy distribution under vehicle acceleration/climbing condition (TMS: Thermal Management System)

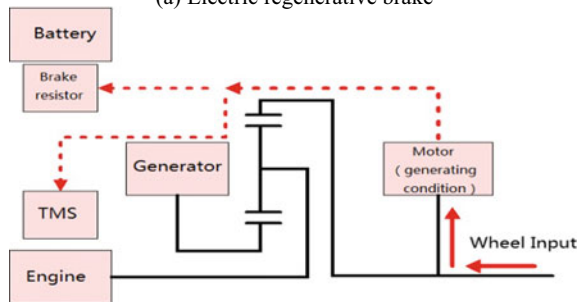


Electric brake condition. The energy distribution of series-parallel hybrid electric vehicle under electric brake condition can be seen in Fig. 8. In the brake process, the inertia energy of the vehicle is transmitted to the motor through the transmission system. The electric motor works in the generating mode. It can provide the electricity needed for the thermal management system as well as charge the battery in order to achieve the recycling of the energy. At the same time, the corresponding motor braking torque can be applied to the drive wheel through the transmission system. In this process, the engine and generator does not work. The vehicle's heat

Fig. 8 The energy distribution under vehicle electric brake condition (TMS: Thermal Management System)



(a) Electric regenerative brake



(b) Electric energy consuming brake

source components include the motor, motor controller and transmission oil. When in the high-intensity braking or continuous braking process, because of the capacity limitation of the energy storage device, the recovered energy cannot be completely absorbed. In this case, the brake resistor is needed to transform the electricity into the heat which can be dissipated by the energy consuming cooling subsystem in order to ensure vehicle safety and reliability.

4 Simulation Model and Results Analysis

A. Simulation model

According to the above analysis, an power delivery energy management strategy is applied in the simulation based on the power demanding. Based on the unique vehicle structure and schematic, a coupled vehicle model which combined the driveline model and the thermal management system model is built in GT suite and applied in the study which is shown in Fig. 9. The energy management strategies are built in details with MATLAB/SIMULINK which is shown partly in Fig. 10. According to the coupled model, the energy flow, energy distribution and heat characteristics of components under different driving condition can be studied in details.

B. Simulation results analysis

A simulation of prototype heavy duty vehicle test driving cycle which can be seen in Fig. 11 is applied in this research. According to the theoretical qualitative analysis and simulation quantitative analysis, power flow and heat flow corresponding relationship

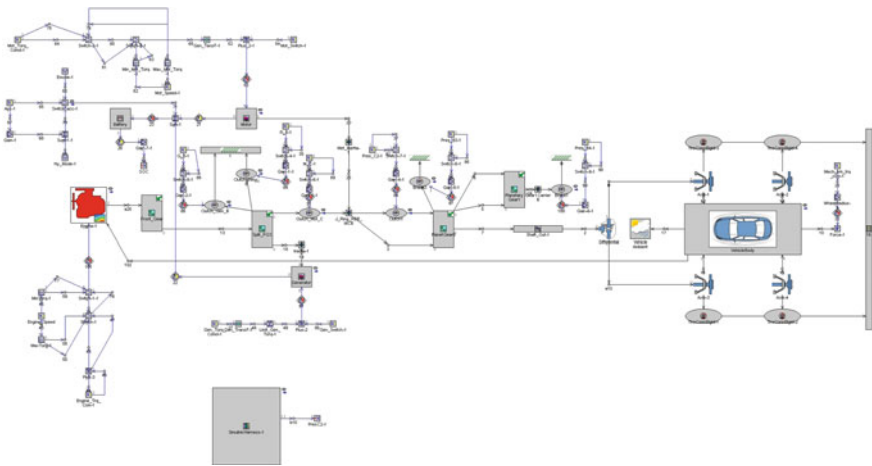


Fig. 9 Coupled vehicle simulation model built in GT suite

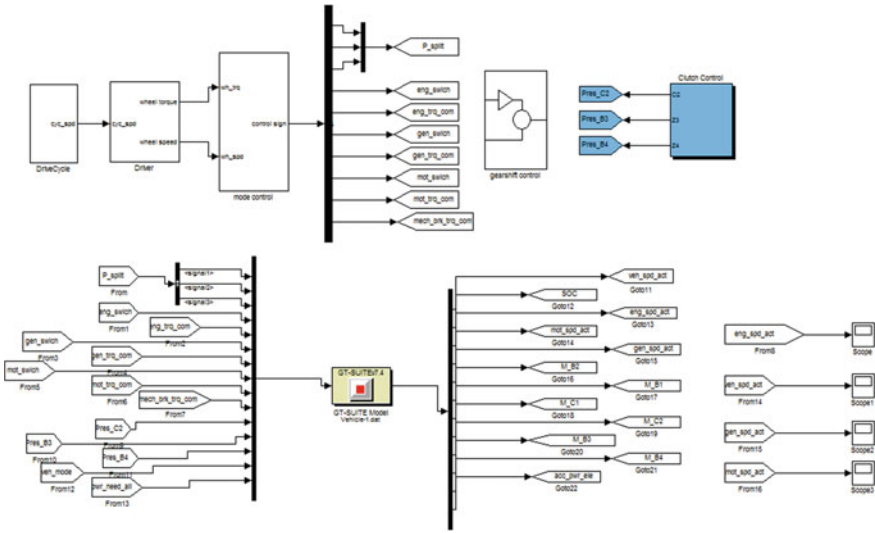


Fig. 10 Whole vehicle control strategies built in MATLAB/SIMULINK

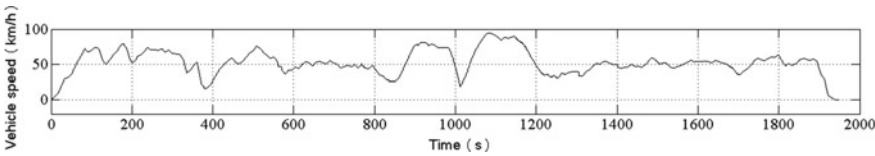


Fig. 11 Heavy duty vehicle test driving cycle

is acquired which can be shown in two aspects. One is vehicle dynamic heat response. The other one is the vehicle power demanding related to thermal management system power consumption. Based on the multi-disciplinary integrated simulation model, some detailed research is conducted to study the dynamic heat response, the power and heat delivery and power consumption under different operating conditions and working modes. Figure 12 shows the engine and electric components temperature histories over the test driving cycle under the rule-based cooling control strategy. Figure 13 shows vehicle thermal management system power consumption over the combined test driving cycle.

The simulation results show that the control strategy of the cooling system can not only satisfy the heat dissipation requirements of the heat source components during the combined driving cycle, but also keep the outlet temperature lower than the limit value. Under this circumstance, the hybrid driving components can work effectively in the better temperature range, and effectively reduce the thermal management system power consumption. Compared with the prototype car using switch cooling control strategy, the power consumption can be decreased by about 30.6% when the ambient temperature is 20°C which be set as the reference ambient temperature.

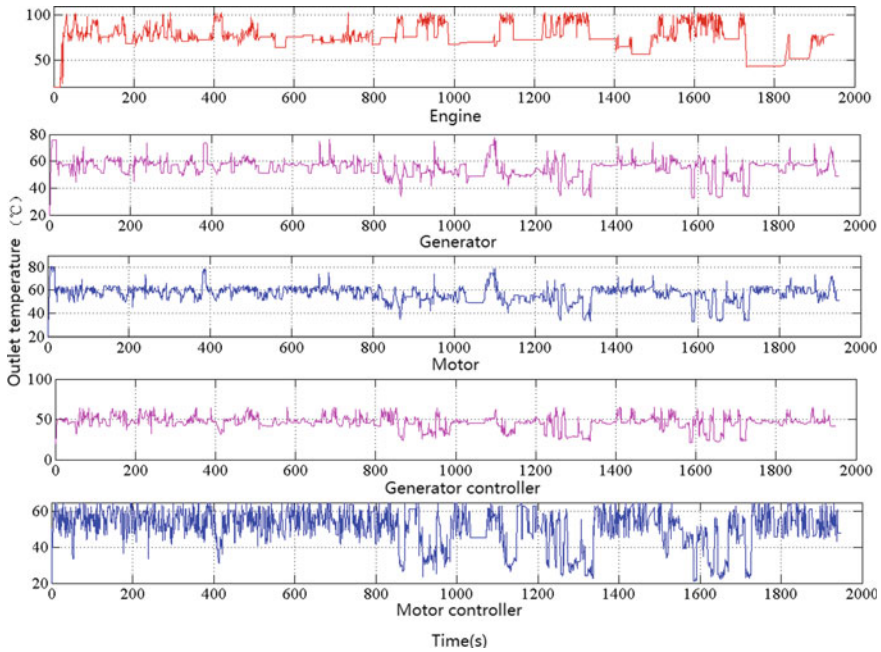


Fig. 12 Temperature histories of engine and electric components under test driving cycle

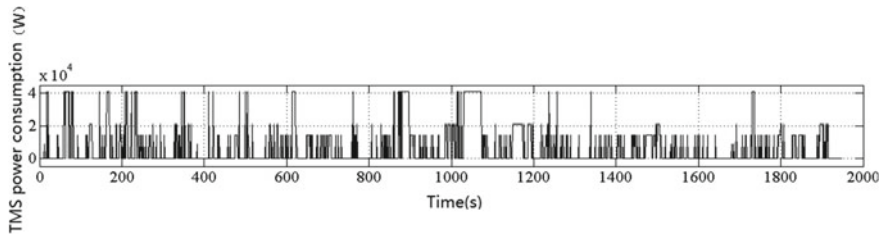


Fig. 13 Thermal management system power consumption under test driving cycle

With the increasing of the ambient temperature, the thermal management power consumption will increase accordingly. When the ambient temperature is 40°C, the thermal management power consumption will be 1.6 times of the reference ambient temperature.

5 Conclusions

In series-parallel hybrid electric vehicle, the relationship between power flow and heat flow is very complex which can influence the vehicle dynamic heat response and

power consumption under different driving condition. The coupled simulation model can solve this complex problem which can be effectively used in the study combined dynamic heat response, the power and heat delivery and power consumption under different operating conditions. This integrated research method combined theoretical qualitative analysis and simulation quantitative analysis is an efficient way in the multi-disciplined complex vehicle system energy management study.

Acknowledgements Fanatical and technical supports from the National Natural Science Foundation of China under Grant 51705480 and China North Vehicle Research Institute are gratefully acknowledged.

References

1. Miro-Padovani T, Colin G, Ketfi-Cherif A, Chamaillard Y (2016) Implementation of an energy management strategy for hybrid electric vehicles including drivability constraint. *IEEE Trans Veh Technol* 65(8):5918–5929
2. Montazeri-Gh M, Mahmoodi-K M (2016) Optimized predictive energy management of plug-in hybrid electric vehicle based on traffic condition. *J Clean Prod* 139:935–948
3. Zhang S, Xiong R, Sun FC (2017) Model predictive control for power management in a plug-in hybrid electric vehicle with a hybrid energy storage system. *Appl Energy* 185:1654–1662
4. Park S, Jung D (2010) Design of vehicle cooling system architecture for a heavy duty series-hybrid electric vehicle using numerical system simulations. *J Eng Gas Turbines Power* 132:092802–092802–092811

Information Extracting and Analysis on Safe Operation of Online Car-hailing Based on Text Mining: The Case of China



Si Yang Liu, Jian Rong, Chen Jing Zhou, Yi Wang, and Zi He Lin

Abstract The recent rise of online car-hailing has successfully promoted the booming of sharing economy in transportation and sets up new mode for intra-city trips. However, it comes with serious risk for passengers, drivers and online car-hailing platforms (OCPs). Few researches have explored the risk of OCPs services despite most research focus on the safe assessment/evaluation system. As a special transportation service enterprise, the research on the safe of OCPs is obviously insufficient. This paper explores the focus of online car Hailing industry-related researches, and build a framework for OCPs services by identifying 8 topics from 883 research papers. Natural language processing method are adopted to achieve the target and specifically, the TF-IDF algorithm and LDA model are applied in the study. The results show that regional policy background provides basic safe situation for OCPs services, and stress the importance of safety improvement from both macro vision and detailed optimization method. Besides, current disputes and dilemma of OCPs safe operation and the legal status and responsibility of OCPs should be paid more attention. In addition, shielded from various risks (e.g., functional risk, time risk physical risk, financial risk) is a key to get the advantage in the competition for market share. Finally, practice and planning of OCPs services in regions is needed to promote the safe operation of OCPs services.

Keywords Online car-hailing · Text Mining · Sharing economy

1 Introduction

As a new product of the sharing economy, the recent rise of online car-hailing has successfully promoted the booming of sharing economy in transportation and sets up

S. Y. Liu · J. Rong · Y. Wang · Z. H. Lin
Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing, China

C. J. Zhou (✉)
Beijing University of Civil Engineering and Architecture, Beijing, China
e-mail: zhouchenjing@139.com

new mode for intra-city trips [3, 22, 24, 31]. In contrast to the traditional taxi service, the online car-hailing service allows passengers to submit trip requests to the drivers nearby through a smartphone with related applications and the system responses immediately with necessary vehicle and driver information (e.g., vehicle location, number of license plate, potential waiting time and possible travel fee). Meanwhile, online rail-hailing drivers accept trip requests when cruising on the road, and pick-up passengers from given locations [3]. The new travel pattern reduces waiting time for both passengers and drivers by matching passenger travel information and vehicle service information effectively and making full use of idle vehicles which improves travel service and sources allocation efficiency [9]. In addition, different levels of vehicles and services are available for users and the charge in different levels of service is completely determined by the market and complete the trade through online electronic transactions. Hence, compared to the single service that traditional taxi service provide, online car-hailing satisfies users with various travel quality requirements and avoids the contradiction of cash transaction.

Due to the high efficiency, reliability and low operating costs, online car-hailing service have been well received by users and drivers, stimulating huge market demand [1, 23, 25, 26]. Several online car-hailing Platforms (OCPs) (also called Online Car-hailing Companies or Transportation Network Companies), such as Uber, Lyft, and Didi, come with a huge number of passenger/driver members and the prosperity market arisen worldwide. Uber is a highly valued OCP that launched in 2010 in San Francisco as the first OCPs to provided city inhabitants with a convenient and reliable door-to-door mobility service, which covers in more than 700 cities currently. It's noticed that Uber's service area has been expanded into more peripheral suburban neighborhood though it widen the market in metropolitan cities [5, 7]. Didi, launched in 2012, is the largest OCP worldwide with 450 million users and 21 million drivers which leads to more than 30 million trips per day in more than 400 Chinese cities [21, 25]. Affected by OCPs services, the ridership of traditional taxi continues to decline in various cities around the world, particularly in Beijing and the USA. In Beijing, with the gradual expansion of OCPs services, the passenger volume of traditional taxi was 14% smaller in 2018 than it was in the 2017, and approximately 50% smaller than in 2012 [2]. Similarly, there are 170,000 daily trips through OCPs, which is almost 12 times than traditional taxis and accounts for at least 9% modal share in San Francisco [25].

However, a large market share and upward tendency of OCPs services come with serious risk for passengers, drivers and OCPs, and the concerns of the safe of OCPs services, such as crime or an unpredictable incident, are increasing worldwide [5, 12, 20]. According to the statistical data from China Justice Big Data Service Platform, the disputes and conflicts between drivers and passengers was 50 and 42% for OCPs services and traditional taxi services, respectively.¹ For example, a woman was sexually assaulted by a Lyft driver in 2017, a stewardess was raped and murdered by the member driver of DiDi en route on the evening of May 5, 2018.² However,

¹ <http://data.court.gov.cn/pages/research.html>.

² https://www.sohu.com/a/231280374_459929.

note that Uber's terms and conditions state that the entire risk of using the service lies on the user-side [10]. OCPs service has not only changed the traditional taxi industry operation mode, but also established a new legal relationship between OCPs drivers, passengers and the platform[12]. The relationship becomes intensive and always be widely discussed by citizens due to the unclear division of accident responsibility when an accident occurs.

We attempt to construct and propose a safety operation management framework for OCPs services, clarifying the safety topics among uncertain. The paper proceeds by first offering brief information on the safety of OCPs services and a review of previous safe operation system analysis approaches in the literature. Next, we present a methodology for extracting the safety-related items for OCPs service. Then, we analysis and discusse results, define the outcome. Finally, the paper concludes with contributions and limitations of the method.

2 Related Work

As part of our research, related work pertaining to all aspects of OCPs services safety was conducted and concentrated on the safe of OCPs services and the practice for ensuring the safe of OCPs service operating.

2.1 *The Safe of OCPs Sevices*

Many researches have attempted to identify safe-related indexes and by establishing the safe assessment system to evaluate the security or promote the improvement of service quality of the OCPs services[6], instead of the safety aspect, hence the efforts to prevent risk are still insufficient. Few researches for safe of OCPs services focus on the physical risk, perceived risk and information security of passengers. In one of the few academic studies, Hong classified the risk that OCPs passenger faced with into the physical risk and the performance risk [10]. The former is defined as the risk of direct damage to personal physical and mental health in the process of OCPs services, which is the most important factor affecting passengers travel model choice [15]. For instance, as mentioned above, a stewardess was raped and murdered by the member driver of DiDi en route in 2018. The performance risk refers to the possibility of feeling buyer's remorse, such as the sense of regret after having made a purchase, due to the discrepancy between the expected value and the actual value. Ma proposes an integrated model to explore how perceptions of risk affects a person's decision to use OCPs services. The result shows that users' perception of physical risk is the most important factor negatively affecting users' trust in drivers and the trust has a positive effect on trust in OCPs [20]. Chaudhry discusses the passenger safety in OCPs services, and believes that mandatory itinerary sharing and uploading,

real-time video monitoring and passenger insurance for passengers can reduce the physical risk of harassment, attack and robbery [5].

In addition to physical risk and perceived risk, the information security of drivers and passengers is also one of the concentration of research [16, 19, 27, 30]. For instance, Li thinks that passengers are easy to cause personal privacy disclosure in location-based services, and hence prefer to use the landmark entities around them to describe their approximate location than their exact location when they ordering OCPs service [18, 27]. Therefore, Li and Wang proposes a location privacy protection scheme based on Minhash algorithm to select a certain range of landmarks near the user's exact location as his location features. Meng also believes that sensitive information may be leaked in the process of OCPs services, and then develops a method of collaborative taxi service with privacy protection assisted by blockchain, which ensures the security of information flow by anonymous verification of passenger and driver information, conditional disclosure of information, creation of records and platform to match passenger and driver pairing [17].

Limited research has focus on the safety research of OCPs, but as one of the transportation service industries, OCPs still have the similar safe problems shared by transportation enterprises or service industry enterprises. It is pointed out that service enterprises are faced with risks of natural disasters, business activities, technological changes and changes in user preferences [11]. Based on questionnaires and interviews, assessed 487 transportation enterprises' risks in Slovak [13]. The results showed that the risk of enterprises in descending order was market risk (26%), financial risk (21%), economic risk (19%), personal risk (11%), operational risk (9%), legal risk (7%), security risk (6%) and other risks (1%). Transportation enterprises pay more attention to the risk of enterprise economy, but ignore the risk of operation safety. Bian and Gong established the safety evaluation system of transportation enterprises from the three dimensions of comprehensive management, driving management and vehicle management. The former thinks that the importance of comprehensive management is the most important, while the latter thinks that the importance of driving management is the highest [3, 8]. However, the risk management of transportation industry is not systematic, such as incomplete consideration of risk content, unclear risk responsibility, serious lack of relevant standard system, etc. It is worth noticing that the OCPs is more different from the traditional taxi industry that OCPs drivers can use private cars to pick up passengers through the platform. The safety of passengers depends on their driving technology and environmental characteristics, and the safety operation risk is relatively higher [4, 5].

2.2 Practice for Ensuring the Safe of OCPs Service

At the practical level, local governments have made detailed and strict regulations on how to reduce passenger safety risks, improve service quality to attract passengers

and ensure travel safety, mainly including the access threshold of drivers and vehicles, the content and quality requirements of transportation services, and the operation and maintenance, management and service capabilities of platform companies. The United States defines the enterprises engaged in OCPs service as TNC (Transportation Network Companies), and states and cities put forward strict restrictions on drivers, operating vehicles and TNCs themselves. In October 2015, the Ministry of road and transport of India issued the guidance on the supervision of online car hailing, which requires that the online car Hailing platform must have sufficient service capacity, and put forward strict access threshold and requirements for operating vehicles and drivers. In addition, the United Kingdom, France and Japan and other countries have similar requirements for OCPs.³

Taking Seattle as an example, Uber and Lyft are the main TNCs recognized at present. The local government strictly supervises them and puts forward strict binding clauses for drivers, operating vehicles and TNC enterprises. In Seattle, OCPs drivers are required to have certain driving qualifications, such as clean background (no criminal discipline, drinking driving, etc.), holding valid driving certificates and receiving the formulated training and education. Meanwhile, the responsibilities of drivers are specified, including payment on online platform, no long-term operation, weekly reporting of operation information, etc. Provide safe and stable driving, keep the interior environment comfortable, clean and sanitary, and help passengers carry luggage. In addition, drivers are also required not to operate illegally or illegally (such as drinking, drug transportation, discarding goods, etc.), driving without using mobile phones, smoking, refusing to carry (guide dog, luggage) without reason, and not to offend or discriminate against others. The illegal behaviours will be severely punished. In terms of vehicles, the operating vehicles are required to have certain operation qualifications, such as registration in transportation network companies, acceptance of the municipal government's inspection, insurance, etc. In addition to meeting the requirements of normal operation companies, TNCs also require to obtain TNC license, provide driver and vehicle information, keep all operation records for 2 years, have a lost and found system, provide services during working hours on weekdays, and assist and implement the local government's requirements for drivers and operating vehicles.⁴

Most research cited above focused on the physical risk, perceived risk and information security of passengers. However, as a special transportation service enterprise, the research on security risk of OCPs is obviously insufficient. At the practical level, local governments have made detailed and strict regulations on how to reduce passenger safety risks, improve service quality to attract passengers and ensure travel safety, including the access threshold of drivers and vehicles, the content and quality requirements of transportation services, and the operation and maintenance, management and service capabilities of platform companies. As the development of computer technology, many advanced algorithms, including text mining, set a new

³ https://www.sohu.com/a/116456379_442411.

⁴ <http://www.seattle.gov/business-regulations/taxis-for-hires-and-tncs/transportation-network-companies>.

path for exploring the safety of OCPs services. Text mining is a process of extracting high-quality information from text or web documents [29]. Hence, inspired by the previous researches, practice and the new method, this paper proposes a safety operation management framework for OCPs services based on natural language processing, constructing and proposing a safe operation management framework for OCPs services.

3 Data and Methodology

3.1 Data Sources

China were selected because many efforts have already been made to steer and regulate smart mobility since the OCPs service first appeared. The supervision and legal regulation of OCPs services, the legal relationship of OCPs, the labor legal relationship related to OCPs drivers and insurance are widely discussed [12]. Academic papers and newspapers are important data sources to understanding related topic, hence in this paper, we used these sources to conduct citation analysis and text mining to extract and analyze the safe operation items of OCPs services. Academic papers and newspapers in this paper collected from CNKI with the key words of online car-hailing and safety in 9th July, 2020, including title, abstract, keywords and year of publish. The left of Fig. 1 illustrates the research numbers for each year while the right of Fig. 1 shows the keyword in this research, representing the research trend in this field. Note that, all paper we collected are write in Chinese, some figures hence present in row language. As shown in Fig. 1, papers have been written since 2015 and reach the climax in 2017. The majority papers focus on the relationships among varies participants, including vehicles, passengers, drivers and government.

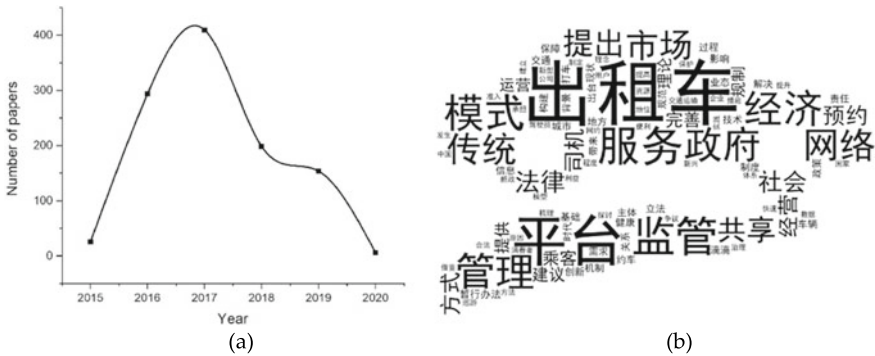


Fig. 1 Research numbers (a) and keywords (b)

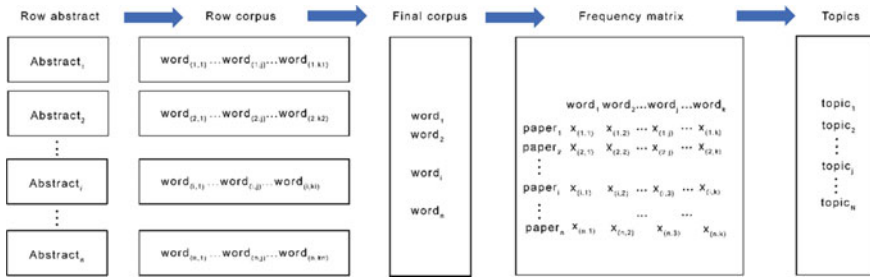


Fig. 2 Working procedures

3.2 Method

For further exploring the safe operating framework for OCPs services, text mining method is adapted to achieve the target. Figure 2 illustrates the working procedures in this paper and we use the abstract in collected paper as basic analysis source.

As shown in Fig. 2, we:

- (1) preprocess all abstracts to get row corpus. In order to achieve the ideal effect of word segmentation and avoid the interference of function words such as “and”, “so”, “hence”, etc., user-defined function word table are selected to delete the word and improve the accuracy of text mining. However, the row corpus could include the words come in most of paper without any features for identifying the papers despite the drop of function words. For example, “suggestion” may appear in all paper and it provide limited information for the extraction. Hence, in the second procedure, we
- (2) use TF-IDF algorithm to extract accurate words that can reflect a given paper’s feature to build the final corpus. TF-IDF is a statistical method to evaluate the importance of a word to a document set or one of the documents in a corpus. The main idea of TFIDF is that if a word or phrase appears frequently in one article and rarely appears in other articles, it is considered that the word or phrase has good classification ability and is suitable for classification. The importance of a word increases in proportion to the number of times it appears in the document, but at the same time it decreases inversely with the frequency of its appearance in the corpus. In a given document, term frequency (TF) refers to the frequency of a given word in the document which can be expressed as follows:

$$TF_{(i,j)} = \frac{c_{(i,j)}}{\sum_{i=1}^n c_{(i,j)}} \tag{1}$$

where $TF_{(i,j)}$ refers to the term frequency for word j in abstract i ; $c_{(i,j)}$ refers to the times of word j appears in abstract i ; n refers to the total number of abstracts.

Inverse document frequency (IDF) is a measure of word universal importance. The IDF of a specific word can be obtained by dividing the total number of files by the number of files containing the word, and then taking the logarithm of 10 as the base to get the quotient:

$$IDF_j = \lg \frac{n}{N(t_j \in d_i)} \quad (2)$$

where IDF_j refers to the inverse document frequency for word j ; $N(t_j \in d_i)$ refers to the document numbers that contains word j .

Then select important words which represents the features of each abstract to form the final corpus after calculating the product of TF and IDF as TF-IDF value to represent the importance of word j in document i . in the next step, we.

3) build frequency matrix based on the word in final corpus and related TF-IDF value for each paper and word. As shown in Fig. 2, $x(i, j)$ refers to the TF-IDF value for word j in paper i . We then.

4) apply late Dirichlet allocation (LDA) model to identify topics and key words based on the frequency matrix. Before applying the LAD model, the similarity test was conducted to find a proper topic number. Cosine similarity (CS) is to measure the difference between two individuals by the cosine value of the angle between two vectors in a vector space. The cosine similarity value is between 0 and 1, and the cosine value is close to 1 which indicates that the more similar the two vectors are. CS is expressed as follows:

$$CS_{i,j} = \frac{word_vec_i \cdot word_vec_j}{|word_vec_i| \times |word_vec_j|} \quad (3)$$

$$CS = \frac{\sum_{i=1}^N \sum_{j=1}^N CS_{i,j}}{N^2} \quad (4)$$

where $word_vec_i$ refers to the word frequency vector of topic i ; $CS_{i,j}$ refers to the cosine similarity of topic i and topic j . N refers to the total topic number.

After training the LDA model, we tag all papers with varies topics and form the items matrix for OCPs safe operating via the key word that identified by LDA. Hence, a framework for OCPs safe operating therefore formed and we discussed the framework in the final part.

4 Results

Table 1 shows the row corpus after the preprocess. In the row corpus, there are some words cannot reflect the feature of the paper’s abstract, hence we selected the words with frequency greater than 10 for the following analysis and calculation.

As mentioned in Sect. 3.2, TF-IDF value are applied to identify feature words for each paper. Table 2 presents the TF-IDF value for each word in row corpus in paper which ID is 0. For further clarifying the words with higher importance, we then select the words with TF-IDF value greater than 0.1 to form the final corpus and build the weighted frequency matrix with TF-IDF value.

For exploring the proper topic numbers, CS values were calculated for topic numbers from 2 to 49 and the result are illustrated in Fig. 3. Therefore, 8 with CS value of 0.149 is chosen as final topic number duo to the CS maintains lower values

Table 1 Word frequency in the row corpus (partial)

Paper ID	Special car	Format	Personalization	Duty	Centre	Subject
0	0	0	0	1	0	0
1	0	0	0	0	0	1
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	1	0	1	1	0
...

Table 2 TF-IDF values for words in paper 0 (partial)

Words	Violation	Jjudgment	Situation	Particularity	Employment	...
TF-IDF	0.351	0.335	0.298	0.272	0.269	...

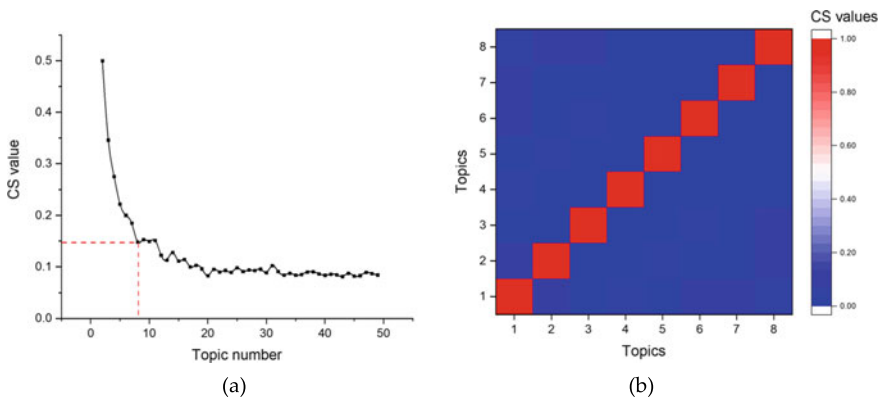


Fig. 3 CS values under different topic number (a) and CS values matrix for 8 topics (b)

Table 3 Topics with its words and probability

Topics	Words				
Topic 1	New policy (0.022)	Transportation (0.020)	Nation (0.019)	Release (0.016)	Opinion (0.015)
Topic 2	Control (0.014)	Administration (0.013)	Measure (0.010)	Monopoly (0.009)	Level (0.009)
Topic 3	Factor (0.014)	Life (0.013)	Model (0.012)	User (0.010)	Improvement (0.009)
Topic 4	Enterprise (0.015)	Reason (0.013)	Risk (0.013)	Online appointment (0.013)	Occur (0.012)
Topic 5	Cause (0.011)	Operation (0.009)	Enterprise (0.009)	Activity (0.008)	Fault (0.008)
Topic 6	Carding (0.009)	Public transportation (0.008)	Game theory (0.007)	Case (0.007)	Model (0.007)
Topic 7	Cognizance (0.013)	Tort (0.011)	Undertake (0.011)	Accident (0.010)	Definition (0.010)
Topic 8	Measures (0.010)	Experience (0.009)	Cooperation (0.008)	Principle (0.008)	Develop 0.007

and change less when topic number is larger than 8. The right chart of Fig. 3 shows the cosine similarity matrix for 8 topics, and all CS values are less than 0.1, indicating the excellent inconsistency for each topic pair.

We finally applied LDA model via Jupyter Notebook to exploring words and related content of 8 topics based on the weighted frequency matrix with TF-IDF value. The first five words with maximum probability to appear in each topic are presented in Table 3.

5 Discussion and Conclusion

As shown in Table 3, 8 topics are identified by applying LDA model and so the key words. In topic 1, key words include new policy, transportation, nation, release and opinion. These words indicate the regional policy background provides basic safe situation for OCPs services. Quickly expansion of OCPs services may explain by the willingness of OCPs to operate outside the law in unregulated markets in some countries or in grey regulatory areas [25]. However, in 2016, the OCPs services was legitimated and restricted in China after the promulgation of the guiding opinions on deepening the reform and promoting the healthy development of the taxi industry, the Interim Measures for the management of online booking taxi business services, and the detailed rules for the management of online booking taxi business services (the new policies). In addition to the policy background, economic situation plays an important role in the safe operation of OCPs services. Indeed, the OCPs services

entered the period of normative adjustment due to the macro-economy slowed down in 2019 [28].

The words in topic 2 stress the safety improvement of the OCPs industry from a macro vision while topic 3 focus on the detailed method and ways to optimization OCPs services security. From the macroscopic view, in the past years, Didi has incomparable advantages in terms of transport capacity scale, and with the expansion of capital, it still dominated the market. Fortunately, new policies were issued to against monopolization in OCPs services and standardize the services, stimulating the healthy competition of OCPs service, hence improving the quality and safety level of the services. For passengers, more choices are available due the healthy market situation. Topic 3 discusses the impact of the safety factors and proposes theoretical model to guide the operation of OCPs on how to improve the safety.

Topic 4 describes the current disputes and dilemma of OCPs safe operation and topic 5 express the legal status and responsibility of OCPs in an accident. First, as mention in Sect. 1, a large market share and upward tendency of OCPs services come with serious risk for passengers, drivers and OCPs. The OCPs provide the services and intergraded the information of drivers and passengers, which makes it lie on the cusp of public opinions [5, 12, 20]. Second, new policies were released to preserve order and discipline of OCPs services and market, but the requirement and supervision for OCPs were serve at the same time. In addition, accurate identification of the legal status of OCPs is a key to solve the safety problem of OCPs services. However, there are many different views on the legal status of OCPs, leading to virous legal responsibilities of OCPs [14].

Topic 6 emphasizes the competition of OCPs services and on traditional transportation field. The OCPs services is well received for its high efficiency, reliability and low operating costs, forming a fierce competitive relationship with public transport [25]. However, various risks (e.g., functional risk, time risk physical risk, financial risk) has a negative effect on intention to continue use of OCPs services [20]. Topic 7 concentrates on the legal relations among virous participants, particularly in an accident. The words of measures, experience, cooperation, principle and develop consist the topic 8 and refer to practice and the future of OCPs services in regions. As mentioned in Sect. 2.2, local governments have made detailed and strict regulations for drivers, OCPs to reduce safety risk. As for the develop trend, the online car hailing industry will develop with integration, business model facing innovation, and operation will become more refined due the stable policy, increasing economy, and therefore the market will present a good development prospect.

In a conclusion, this paper explores the focus of online car Hailing industry research through the natural language processing method, and a total of 8 safety aspects are identified. However, this paper focus on the abstract of each paper, assuming that all abstracts are well written and provide key information of the related paper, which may ignore some key content in the full text. In addition, the LDA model used in this paper sets up a new path for exploring the safety of OCPs service, but the model based on the bag-of-words model which neglects the relationships between words.

References

1. Acheampong RA, Siiba A, Okyere DK, Tuffour JP (2020) Mobility-on-demand: an empirical study of internet-based ride-hailing adoption factors, travel characteristics and mode substitution effects. *Transp Res Part C Emerging Technol* 115(March):102638. <https://doi.org/10.1016/j.trc.2020.102638>
2. Beijing Transportation Research Center (2019) Beijing Transport Annual Report: 2. <http://www.bjtrc.org.cn/List/index/cid/7.html>
3. Bian H, Wang Y (2009). The safety evaluation on road passenger transportation enterprises based on modified AHP method. In: 2009 sixth international conference on fuzzy systems and knowledge discovery, IEEE, 413–16. <http://ieeexplore.ieee.org/document/5359198/>
4. Cetin T, Deakin E (2019) Regulation of taxis and the rise of ridesharing. *Transport Policy* 76: 149–58. <https://doi.org/10.1016/j.tranpol.2017.09.002>
5. Chaudhry B, Yasar A-U-H, El-Amine S, Shakshuki E (2018) Passenger safety in ride-sharing services. *Procedia Comput Sci* 130:1044–1050. <https://doi.org/10.1016/j.procs.2018.04.146>
6. Epifanov V, Obshvalkin M, Lukonkina K (2018) Management of quality and security level of transportation in the system of regular passenger motor transport. *Transp Res Procedia* 36:141–148. <https://doi.org/10.1016/j.trpro.2018.12.056>
7. Gehrke SR (2020) Uber service area expansion in three major American cities. *J Transp Geogr* 86:102752. <https://doi.org/10.1016/j.jtrangeo.2020.102752>
8. Gong H (2011) A study on comprehensive evaluation of road transportation safety based on grey relational grade. In: ICTIS 2011. American Society of Civil Engineers, Reston, VA. 770–77. <http://ascelibrary.org/doi/10.1061/41177%28415%2998>
9. Hawlitschek F, Teubner T, Gimpel H (2018) Consumer motives for peer-to-peer sharing. *J Clean Prod* 204:144–157. <https://doi.org/10.1016/j.jclepro.2018.08.326>
10. Hong JH, Kim BC, Park KS (2019) Optimal risk management for the sharing economy with stranger danger and service quality. *Eur J Oper Res* 279(3):1024–1035. <https://doi.org/10.1016/j.ejor.2019.06.020>
11. Hu B, Feng Y (2017) Optimization and coordination of supply chain with revenue sharing contracts and service requirement under supply and demand uncertainty. *Int J Prod Econ* 183:185–93. <http://dx.doi.org/10.1016/j.ijpe.2016.11.002>
12. Huang D, Zhang Y (2020) Research on Tort liability of network car platform under social vehicle mode: based on judicial case studies of online car infringement. *J Chongqing Univ Posts Telecommun (Soc Sci Edn)* 32(2): 44–53
13. Hudáková M, Masár M (2018) The assessment of key business risks for SMEs in Slovakia and their comparison with other EU countries. *Entrepren Bus Econ Rev* 6(4): 145–60. <https://eber.uek.krakow.pl/index.php/eber/article/view/472>
14. Jiahui GUO (2020) On the legal responsibility and supervision improvement of ride-hailing platform. *J Henan Polytech Univ (Soc Sci)* 21(5)
15. Kaplan LB, Szybillo GJ, Jacoby J (1974) Components of perceived risk in product purchase: a cross-validation. *J Appl Psychol* 59(3): 287–91. <http://content.apa.org/journals/apl/59/3/287>
16. Khazbak Y, Fan J, Zhu S, Cao G (2018) Preserving location privacy in ride-hailing service. In: 2018 IEEE conference on communications and network security (CNS), IEEE, 1–9. <https://ieeexplore.ieee.org/document/8433221/>
17. Li M, Zhu L, Lin X (2019) Coride: a privacy-preserving collaborative-ride hailing service using blockchain-assisted vehicular fog computing. In: Lecture notes of the institute for computer sciences, social-informatics and telecommunications engineering. LNICST, 408–22. http://link.springer.com/10.1007/978-3-030-37231-6_24
18. Li M, Wang L (2018) Privacy preservation of location information based on MinHash algorithm in online ride-hailing services. In: 2018 sixth international conference on advanced cloud and big data (CBD) IEEE, 257–62. <https://ieeexplore.ieee.org/document/8530849/>
19. Luo Y, Jia X, Fu S, Xu M (2019) PRide: privacy-preserving ride matching over road networks for online ride-hailing service. *IEEE Trans Inf Forensics Secur* 14(7):1791–1802

20. Ma L, Zhang X, Ding X, Wang G (2019) Risk perception and intention to discontinue use of ride-hailing services in China: Taking the example of DiDi Chuxing. *Transp Res F Traffic Psychol Behav* 66:459–470. <https://doi.org/10.1016/j.trf.2019.09.021>
21. Ma S, Long Y (2020) Functional urban area delineations of cities on the Chinese mainland using massive didi ride-hailing records. *Cities* 97:102532. <https://doi.org/10.1016/j.cities.2019.102532>
22. Moscholidou I, Pangbourne K (2019) A preliminary assessment of regulatory efforts to steer smart mobility in London and Seattle. *Transport Policy*. <https://doi.org/10.1016/j.tranpol.2019.10.015>
23. Pham A et al (2017) Privateride: a privacy-enhanced ride-hailing service. *Proc Privacy Enhancing Technol* 2017(2):38–56
24. Sopjani L, Stier JJ, Mia H, Sofia R (2020) Shared mobility services versus private car: implications of changes in everyday life. *J Cleaner Product* 259
25. Tirachini A (2019) Transportation ride-hailing, travel behaviour and sustainable mobility: an international review. Springer, USA. <https://doi.org/10.1007/s11116-019-10070-2>
26. Tirachini A, del Río M (2019) Ride-hailing in santiago de chile: users' characterisation and effects on travel behaviour. *Transport Policy* 82: 46–57. <https://doi.org/10.1016/j.tranpol.2019.07.008>
27. Wang F et al (2018) Efficient and privacy-preserving dynamic spatial query scheme for ride-hailing services. *IEEE Trans Veh Technol* 67(11): 11084–97. <https://ieeexplore.ieee.org/document/8454748/>
28. Wei ES (2020) Current situation and prospect of online car hailing market *Automob Parts* 8:34–36
29. Xie X et al (2019) A novel text mining approach for scholar information extraction from web content in Chinese. *Futur Gener Comput Syst* 111:859–872
30. Yu H et al (2019) Lpride: lightweight and privacy-preserving ride matching over road networks in online ride hailing systems. *IEEE Trans Veh Technol* 68(11): 10418–28. <https://ieeexplore.ieee.org/document/8842607/>
31. Zuo W, Zhu W, Chen S, He X (2019) Service quality management of online car-hailing based on PCN in the sharing economy. *Electron Commer Res Appl* 34:100827. <https://doi.org/10.1016/j.elerap.2019.100827>

The Method of Division and Classification Quantification for Snow and Ice Removal on Airport Pavement



Ya-ping Zhang, Yu Zhang, Xiao-qing Xing, Ye Chen, Nan Yang,
and Jian Mao

Abstract In the northern cities of China, the winter snow and ice on the airport pavement seriously affects the airport operation efficiency. It is necessary to study the area division and classification quantification of airport pavement snow and ice removal operation in order to scientifically dispatch the equipment group of airport pavement for snow clearance and de-icing. Based on the analysis of the importance of different airport pavements, the airport pavement is divided into different snow removal areas, and the initial priority quantification model is proposed. The gray difference network method is used to predict the operation time of the airport pavement snow removal. Then, the dynamic priority model of the airport pavement is established with the initial priority as the input, and the dynamic status of the airport pavement priority in each stage is generated, and output dynamic priority matrix. Finally, the process of snow and ice removal in a large hub airport is simulated, and the simulation results show that the proposed method is feasible.

Keywords Air transportation · Snow and ice removal on airport pavement · Hub airport · Division · Classification

Y. Zhang (✉) · Y. Zhang · Y. Chen · N. Yang
School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin
150090, China
e-mail: zxlt0905@163.com

Y. Zhang
Beijing E-Hualu Information Technology Co. Ltd, Beijing 100043, China

X. Xing
School of Aeronautical Engineering, Civil Aviation Flight University of China, Guanghan
618307, China

J. Mao
Chengdu Civil Aviation Information Technology Co. Ltd, Chengdu 611430, China

1 Introduction

The snow and ice removal of airport pavement and the scheduling and control of equipment group are important parts of airport pavement management in cold regions in winter, and are also the hot issues of airport ground operation support theory research.

At present, the domestic and foreign research on airport snow removal mainly focuses on the deicing technology. Greg White et al. [1] proposed a prediction model based on snow age, historical temperature and intensity for the polar runway in Antarctica, so as to provide support for the design of polar airport ice and snow runway, which provides another perspective experience for our winter airport snow removal operation. We can transform the snow removal operation into snow leveling operation, and reduce the ice and snow pressure on the runway of the winter airport. To ensure that it meets the airworthiness conditions, so as to replace the ice and snow operation. However, the efficiency of ice and snow compaction and snow removal of equipment group is higher or lower, and what conditions are suitable for snow and ice compaction operation need further discussion. Sajed Sadati et al. [2] evaluated the power demand of the airport conductive concrete pavement, and studied whether the self heating pavement technology can replace the traditional pavement deicing operation. Based on the multi commodity network flow model, Perrier et al. [3] established a mathematical optimization method of road network partition snow removal operation, and proposed two heuristic algorithms to solve the problem. Wilke et al. [4] proposed the overall risk assessment framework of airport ground cooperative operation in view of the fragmentary and unconformity of airport ground operation, and established the process model of airport ground operation based on multi-source data, which provided experience and method for the study of airport snow removal operation. LaBelle et al. [5] established a partition snow removal algorithm based on geographic information system, providing decision support for users to use different snow removal departments. Ange'lica Salazar Aguilar et al. [6] proposed an adaptive neighborhood search algorithm based on mixed integer model under the condition that more than two pavements and snowplows are in the same direction and asynchronous.

Under the premise of uncertain airport ground support equipment operation time, Xu et al. [7] fitted the operation time with normal distribution, evaluated the operation time fluctuation of different service items, established the airport ground equipment scheduling model, and finally solved it by genetic algorithm. They use this method to shorten the operation time of the airport operation equipment group and reduce the waiting time of the aircraft. Nie [8] studied the dispatching data transmission method of airport vehicles based on GSM network and image processing technology. After testing, he concluded that the transmission system designed has reliable positioning accuracy, stable image transmission, and reduced acquisition time. Wang [9] obtained the new route planning results for Airport special vehicles by combining multi-agent system with real-time scheduling algorithm. Finally, the model and algorithm developed by Wang Fang were compared with the manual scheduling method,

which proved the superiority of the new algorithm. Yang [10] designed five modules, including aircraft icing prediction, deicing equipment management, central database, display terminal and monitoring terminal, based on GIS, to realize the tripartite information exchange among airport operation management organization, airlines and deicing company, and improve the efficiency and safety of aircraft deicing scheduling. Ren [11] studied the distributed solution method and centralized solution method of flight ground cooperative support. In the centralized strategy, the model of refueling service and boarding service is constructed and solved by genetic algorithm. In the distributed strategy, taking the aircraft refueling service as an example and minimizing the overall service time as the optimization objective, the airport refueling truck scheduling model is constructed and solved by lingo model. Both methods realize the coordinated scheduling of ground support. Xu [12] divides the airport pavement into three areas: taxiway, runway and apron. By means of path optimization and queuing theory, the collaborative scheduling of airport surface resources is studied, and the airport surface joint scheduling model is established. The research on aircraft deicing scheduling can also provide reference for airport pavement deicing scheduling. Chen [13] studied the apron allocation in the process of centralized airport deicing, and established a scheduling model based on queuing theory, which reduced the total time consumption of aircraft in the airport deicing system. Li et al. [14] designed a parallel scheduling algorithm for aircraft deicing, and solved their goal of maximizing the number of deicing aircraft. The operation results were better than the traditional first come first served algorithm. Xing et al. [15] established the queuing model and scheduling model of airport multi apron deicing based on queuing theory and game theory. Through simulation, it is found that the established model can reduce flight delay and aircraft deicing time. Chen [16] designed the operation monitoring system of aircraft deicing process, predicted the airport deicing parameters (snow type, nozzle number, aircraft type), and established the queuing theory model of aircraft deicing. Chen et al. [17] established a nozzle demand prediction model in the process of aircraft deicing with the goal of minimizing deicing cost and flight delay, and optimized the parameters of de-icing equipment.

In view of the above analysis, there is a lack of research on snow and ice removal methods of airport pavement. Therefore, this paper studies the priority division of pavement zoning process for airport pavement deicing scheduling.

2 Partition and Classification Model

Just as traffic flows in different directions have different rights of way in road traffic, the airport pavement deicing operation area should also have priority. In this paper, the airport pavement operation area is divided, and then the entropy weight method is used to establish the initial priority model.

2.1 Division of snow and ice removal on Airport Pavement

The reflection spectrum of the same kind of ground feature is similar, so the same ground feature (pavement) must be “homogeneous”. Therefore, we can classify the remote sensing images and get a preliminary partition. In this paper, K-means clustering is used to partition the airport pavement. (as shown in Fig. 1).

In the process of departure, the aircraft is pushed out from the apron to the taxiway, and then to the runway through the contact lane to wait for takeoff. On the contrary, it lands on the runway, then passes through the connecting way, taxiway, and finally gets off at the apron. Therefore, the airport pavement system is divided into four operation areas: apron, taxiway, connecting runway and runway. The fast and orderly de icing operation in the four operation areas of the airport is the precondition to ensure the normal arrival and departure of the Airport Flights in winter. The process of aircraft entering the port is shown in Fig. 2, and the departure process is the reverse process.

Therefore, the deicing area is simplified into four operation areas: runway, taxiway, connecting way and apron (Table 1).

Fig. 1 k-means clustering map of airport pavement

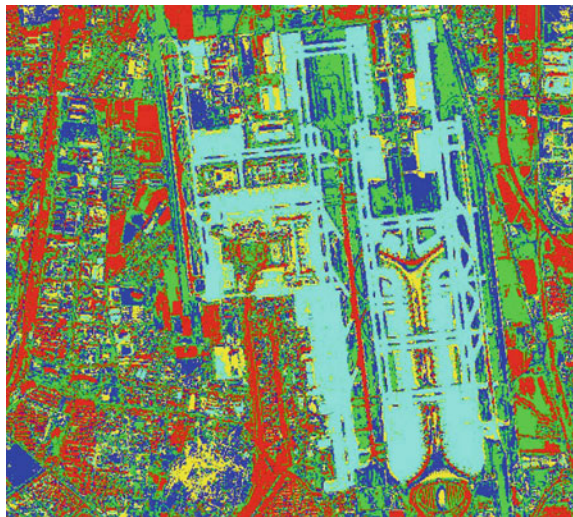


Fig.2 Optimal discount simulating the real order

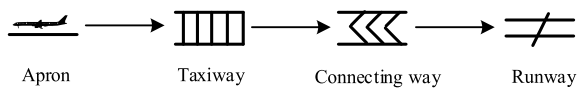


Table 1 Division of snow and ice removal area

Division of snow and ice removal operation	Regional scope
I area	Rrunway
II area	Ttaxiway
III area	Cconnecting way
IV area	Aapron

2.2 Initial priority model

Priority division plays an important role in scientific and orderly de icing of airport pavement. In this paper, a priority model based on pavement thickness, distance and area is established to measure the priority of different pavement by the weight ratio of pavement system.

$$P_i = INT \left[\left(\frac{Z_i - Z_{min}}{Z_{max} - Z_{min}} \right) \times 10 \right] \tag{1}$$

where: P_i is the priority of the i -th deicing operation area. There are four operation zones in this paper, which are apron, runway, connecting way and taxiway. INT means round up. Z_i is the weight of pavement system confirmed by entropy weight method. Z_{max} and Z_{min} represent the maximum and minimum values of a set of weights. The calculated priority of 0 is 1.

In this paper, a weight determination method of pavement system based on information entropy is proposed:

Step 1: data standardization. The data of different dimensions and scales are standardized to facilitate the comparison between data. There are k indexes X_1, X_2, \dots, X_n . The standardized index values are Y_1, Y_2, \dots, Y_n .

$$\begin{aligned}
 X_i &= \{ x_1, x_2, \dots, x_n \} \\
 Y_i &= \{ y_1, y_2, \dots, y_n \} \\
 y_i &= \frac{x_i - \min(X_i)}{\max(X_i) - \min(X_i)} \tag{2}
 \end{aligned}$$

Step2: Calculate the information entropy of each data index. The information entropy of data is defined as the following formula (3).

$$E_j = -[\ln(n)]^{-1} \times \sum_{i=1}^n (p_{ij} \times \ln p_{ij}) \tag{3}$$

Here, $p_{ij} = Y_{ij} / \sum_{i=1}^n Y_{ij}$. In particular, if $p_{ij} = 0$, So, $\lim_{p_{ij} \rightarrow 0} p_{ij} \ln p_{ij} = 0$.

Step3: The weight is determined according to the information entropy of each index. The information entropy of each index is calculated by formula (3), and then the weight of each index is determined according to formula (4).

$$W_i = \frac{1 - E_i}{k - \sum E_i} \quad (i = 1, 2, \dots, k) \quad (4)$$

Step4: According to formula (5), the final numerical evaluation of each different pavement is calculated. I in the formula represents I different pavement, and there are four in this paper. n is the number of evaluation indexes, which is 3 in this paper.

$$Z_i = \sum_{i=1}^n X_{li} W_i \quad (5)$$

Step5: The weight of each pavement priority model parameter is determined by the following formula.

$$R_i = Z_i / \sum_{i=1}^n Z_i \quad (6)$$

3 Dynamic priority algorithm for snow removal area

In this paper, the priority idea of resource scheduling of computer operating system in processing multiple processes is used to complete the dynamic priority allocation of ice and snow operation area, and the dynamic priority algorithm of airport pavement deicing operation area is proposed.

3.1 Grey Differential Network

In this paper, the prediction time is divided into several time slices, which are used to determine the dynamic priority of the work area in the priority algorithm. At the same time, aiming at the shortcoming that grey system analysis can not model the multi input index data, the grey difference network is trained with the sample data. When a new input index group is input, the difference is made between the grey difference network and the prediction index group. The correction value is used to modify the final prediction data, and finally achieve the purpose of data prediction.

Table 2 Varied method of

Difference method	Difference formula
Method 1	$C_{i,G} = x_{r1,G} + \beta^*(x_{r2,G}-x_{r3,G})$
Method 2	$C_{i,G} = x_{r1,G} + \beta^*(x_{r2,G}-x_{r3,G}) + \beta^*(x_{r4,G}-x_{r5,G})$
Method 3	$C_{i,G} = x_{i,G} + \beta^*(x_{best,G}-x_{i,G}) + \beta^*(x_{r1,G}-x_{r2,G})$
Method 4	$C_{i,G} = x_{best,G} + \beta^*(x_{r1,G}-x_{r2,G})$
Method 5	$C_{i,G} = x_{i,G} + \beta^*(x_{best,G}-x_{r1,G}) + \beta^*(x_{r2,G}-x_{r3,G})$
Method 6	$C_{i,G} = x_{i,G} + \alpha^*(x_{r1,G}-x_{r2,G}) + \beta^*(x_{r2,G}-x_{r3,G})$

The common difference formula of grey difference network proposed in this paper can be selected according to Table 2. Where $C_{i,G}$ are the difference vectors in the graph model G, and $x_{ri,G}$ are the i -th input vector. α - β is the modified rule coefficient, $x_{best,G}$ are the best data in the sequence data (according to the prediction purpose, it can be the maximum, minimum and median data).

In order to simplify the model, the difference formula used in this paper is as follows: formula (7):

$$C_{i,G} = \beta * (x_{pr,G} - x_{s,G}) \tag{7}$$

where β is 1, $x_{pr,G}$ are predicted data, $x_{s,G}$ are real data.

The grey difference network (Fig. 3) proposed in this paper is a directed graph with single column propagation from top to bottom. The digraph in this paper refers to the network graph model with at least one directed triple (V_1, V_2, F_{12}), where F_{12} is a transfer function, which makes V_1 vertex propagate to vertex V_2 through transfer function F_{12} . Each node of grey differential network model represents a job time. The transfer function of the grey difference network (Fig. 3) proposed in this paper is the GM (1,1) model of the grey system analysis process. The systematic error of the prediction system can be eliminated through the grey difference network.

In order to avoid the single series data prediction too discrete, the input data are arranged according to the snow condition before using the grey difference network to predict the data.

Step 1: input prediction sample data.

A total of 3 columns of data are input. The three columns of input data are the number of snow blowing vehicles, snow pushers and snow sweepers under certain snow conditions. The last column of T_{co} data of snow removal operation time is generated according to the equipment configuration.

Step 2: grey system analysis process.

According to the time data T_{co} of the first n lines of snow removal operation, the GM (1,1) model of grey system analysis is used to generate the prediction time T_{pr} of M line under the configuration of future m lines.

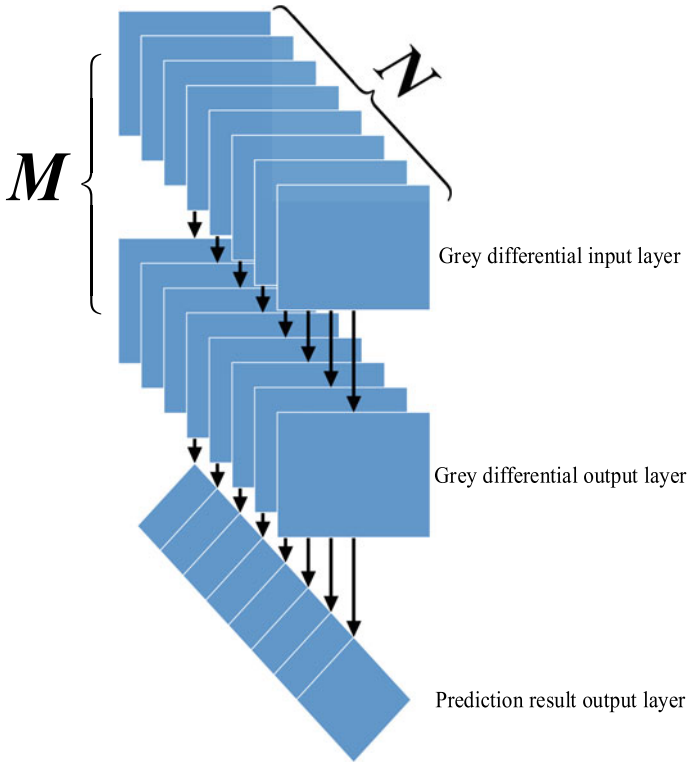


Fig.3 Grey difference network

Step 3: differential correction process.

According to the time T_{co} calculated by m-row configuration, the predicted time TPR is corrected by difference. According to the average difference results of unit equipment configuration in $n + 1$ to $m-1$, the T_{pr} correction of the m-th row snow removal time is allocated. At the same time, based on T_{co} , the operation time T_{pr} was selected.

Step 4: input equipment configuration scheme.

Input the configuration scheme of equipment group A, B, C, D, e and F, and repeat steps 2 and 3.

Step5: output prediction results.

In order to allocate the priority with interval time window, the prediction result should be too large, and the unreasonable data should still be selected based on T_{co} .

3.2 Dynamic Priority Algorithm

The priority of work area is divided into two types: initial priority and dynamic priority. The initial priority is determined in the division of the airport pavement deicing operations. The dynamic priority is estimated according to the average operation capacity of the equipment group and the amount of snow. A dynamic priority algorithm (Fig. 4) is proposed to generate the dynamic priority of work area.

In this paper, the concept of WBC (workspace control block) is proposed. The structure includes: operation area identifier (Name), operation area status (Super), operation area priority (E time), expected deicing time (S time), deicing time (S time).

The operation area identifier (Name) is the unique identification character of the operation area, which is used to distinguish different operation areas. It is represented by integer (int) in the implementation code of this paper. In the actual process, it can also be represented by character type (char).

The process state of priority allocation of airport pavement deicing is not as complex as that of operating system. Therefore, the deicing operation is divided into two states: "end (E)" and "ready (R)".

The purpose of this algorithm is to generate the dynamic priority (Super) of the deicing operation area according to the initial priority generated by the initial pavement partition. Firstly, according to the initial priority, the operation area is arranged into a queue, and the pavement with higher priority is deicing. Then according to the operation area priority (E time), expected deicing time (S time), remaining snow removal time (R Time) and other information to dynamically adjust the priority of the work area. At the same time, set a certain time slice of snow and ice removal, and the time of each time slice is T_{sjp} (10 min, 15 min, 20 min). When the time slice arrives, the priority of the first operation area is reduced by 1, and the remaining operation time is the estimated operation time minus the time slice time. When the priority of the job area in the queue is the same, priority is given to the non running job area.

The priority of deicing operation area is divided into four levels, which is represented by a two-dimensional matrix $P_{n \times m}$, which represents n operation areas (matrix rows) and m equipment groups (matrix columns). The initial priority of runway, taxiway, connecting lane and apron is 1, 1, 4 and 10 respectively. According to the grey difference network, the operation time of runway, taxiway, connecting way and apron is respectively T_r, T_s, T_c, T_i (Table 3).

T_{sjp} is subtracted every time the algorithm runs. When the remaining time is 0, the job area ends the job. During the running process of the algorithm, there are many ending jobs in the operation area. Assuming that the last finished operation area is 3 (connecting way), the last run is m times, and the priority is *, then the state of algorithm m times is shown in Table 4

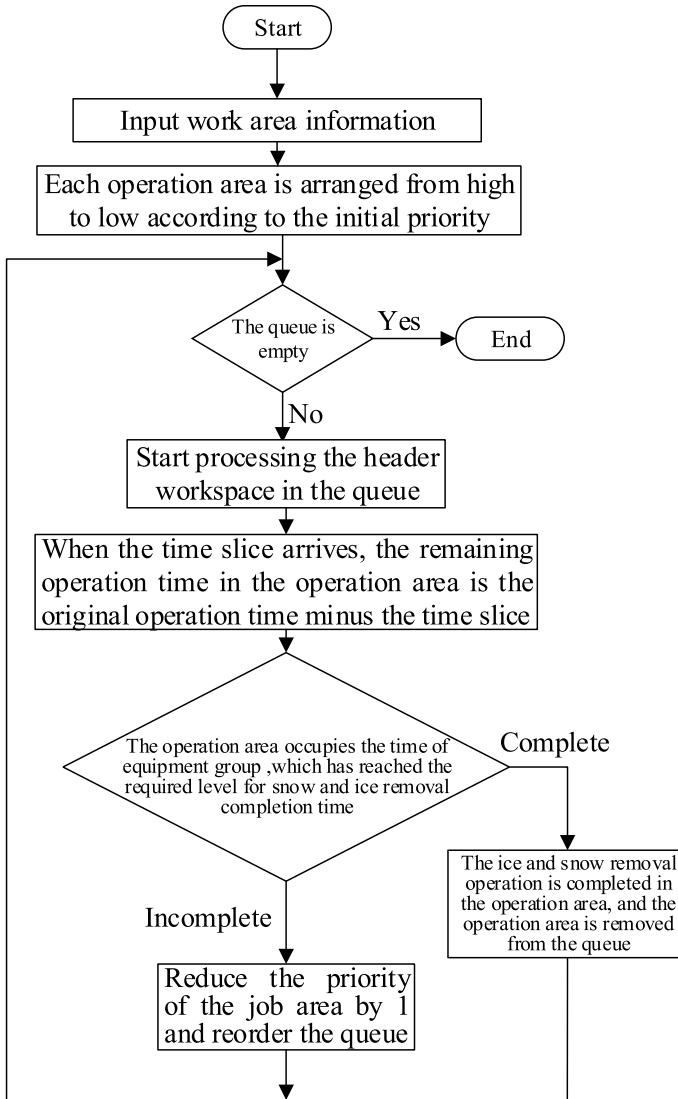


Fig. 4 Dynamic priority algorithm

Table 3 Dynamic priority algorithm runs for the first time

Name	Status	Super	E_time	R_time
1 (runway)	R	1	T_r	T_r
2 (Taxiway)	E	1	T_s	T_s
3 (connecting way)	E	4	T_c	T_c
4 (apron)	E	10	T_t	T_t

Table 4 Dynamic priority algorithm runs for m times

Name	Status	Super	E_time	R_time
3(connecting way)	R	*	T_c	0

4 Simulation Analysis

Taking the second runway of a domestic hub airport as an example, the evaluation indexes of area, distance and thickness of four operation areas are established. According to the priority model proposed in this paper, the information entropy weighting method is used to determine the initial priority of pavement zoning. X_1 , X_2 , X_3 in the table indicate the thickness, distance and area of various pavement respectively. The area can be used to measure the time and distance of snow and ice removal, and can be used to measure the path of equipment group. Because the geometric layout of different airports is similar, the evaluation index has certain universality. The evaluation index is shown in Table 5.

According to the arithmetic mean value of the highest value of various pavement thickness specifications for pavement thickness zoning in *code for design of asphalt pavement of Civil Airport (MH / T5010-2017)*, the arithmetic average thickness of runway and taxiway is 0.77 and 0.75 respectively, and the average thickness of apron is 0.8. The thickness of tie road can not be calculated according to the table of pavement structure thickness in the specification, which is assigned as 0.85. At the same time, the average speed of the aircraft on the runway is significantly higher than that on the taxiway. Therefore, the thickness of runway and taxiway is rounded down and up to 0.7 and 0.8 respectively. The average thickness of the apron is calculated as 0.8, which is taken as 0.9 upward. The vertical distance between each pavement system (runway, taxiway, connecting way, apron) and T3 terminal center is 760 m, 509 m, 666 m and 230 m respectively. The second runway of the airport is runway, and the calculated area is 228000 square meters. There are 8 parallelograms in total, with a height of 150 m and a bottom of 100 m. The corridor rectangle is 20 m wide and 150 m long. There are 19 rectangles in total, including a few arc-shaped ones, which are treated as rectangles. Similarly, the area of taxiway and apron are 177,000 square meters and 1,262,545 square meters respectively. Since the second runway can serve the aircraft launched from the near and far stands of terminal T2 and T3

Table 5 Evaluation index of airport pavement

Pavement	Index		
	X_1	X_2	X_3
runway	0.70	760	228,000
Taxiway	0.80	509	177,000
connecting way	0.85	666	513,000
apron	0.90	230	1,262,545

Table 6 Normalization of evaluation indexes

Pavement	Index		
	X_1	X_2	X_3
runway	0.00	1.00	0.05
Taxiway	0.50	0.53	0.00
connecting way	0.75	0.82	0.31
apron	1.00	0.00	1.00

respectively, the area close to terminal T2 is calculated as half. Table 6 shows the normalized data of evaluation indexes.

According to formula 1–6, the index information entropy E_1, E_2 and E_3 can be calculated as 0.96, 0.96 and 0.69 respectively; the index weights W_1, W_2 and W_3 are 0.1, 0.1 and 0.79; the numerical evaluation indexes Z_1, Z_2, Z_3 and Z_4 are 0.14, 0.13, 0.40 and 0.89 respectively; the model parameter weights R_1, R_2, R_3 and R_4 are 0.09, 0.08, 0.26 and 0.57 respectively. Therefore, according to the initial priority model, the initial priority of runway, taxiway, liaison lane and apron is taken as 1, 1, 4 and 10 respectively, and the queue head in dynamic priority algorithm with small priority value is processed first.

According to Table 7, the composition and operation capacity of the equipment group are taken as Table 7, and the operation time of the hub airport under moderate snow conditions is predicted. The airport terminal covers an area of 1.41 million square meters, the total area of runway is 616 thousand square meters, the area of taxiway is about 820,000 square meters, the area of connecting road is about 930,000 square meters, and the total area of apron is about 4.81 million square meters. If the degree of snow is 40 mm, the snow cover of runway is 24640 m³, that of taxiway is 32800 m³, that of connecting runway is 37200 m³, and that of apron is 192400 m³. According to the operation capacity of equipment group and snow on runway area, T_{co} (unit: minute) can be calculated, as shown in Table 8.

In order to ensure that the scheduling of the airport surface snow removal equipment group will not be chaotic, we always hope that the prediction time is larger. Therefore, this paper leaves a spare time window. Considering that the snow throwing vehicles and snow blowing vehicles have the greatest impact on the operation time, the snow removal operation time is sorted from large to small according to the sum

Table 7 Operation capability of equipment group

Equipment group	Snowblower	Snowmobile	Snowthrower	Operational capability (m ³ /h)
A	2	1	4	28,965
B	2	0	3	23,480
C	3	2	2	17,029
D	1	1	5	31,770
E	0	0	2	11,856
F	4	1	6	36,423

Table 8 Equipment group - operation time of ice and snow removal on road surface

Equipment group	Pavement			
	runway	Taxiway	connecting way	apron
A	51	67	77	398
B	62	83	95	491
C	86	115	131	677
D	46	61	70	363
E	124	165	188	973
F	40	54	61	316

of the number of snow throwing vehicles and snow blowing vehicles. The greater the operation capacity is, the smaller the operation time is, that is, the sequence is from small to large according to the operation time T_{co} (F, D, A, B, C, E). At the same time, the difference between the real value and the predicted value is made to ensure the forward correction. Since the scheduling scheme has not yet taken shape, the arithmetic mean value of each group of differential forward correction data is taken as the priority input of the airport runway, that is, the input predicted operation time of runway, taxiway, liaison lane and apron is 72, 95, 109 and 565 min respectively (as shown in Fig. 5).

In fact, the dynamic priority scheduling algorithm is a method to operate a single pavement at a time. According to the initial priority of the pavement determined in the previous section and the expected snow removal operation time, the input indexes

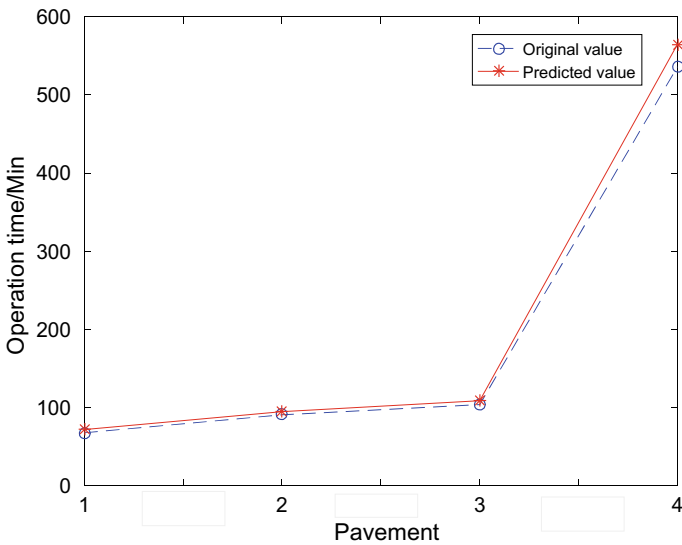


Fig.5 Grey difference network prediction result chart

Table 9 Input index

NAME	initial priority	Expected operation time of snow and ice removal
1	1	72
2	1	95
3	4	109
4	10	565

of the dynamic priority scheduling algorithm are shown in Table 9 (1, 2, 3 and 4 of the name respectively represent the runway, taxiway, connecting way and apron).

The time slice set in this paper is 10, that is, every time the ice and snow operation is run, the expected time will be reduced by 10. Finally, the average priority is taken to generate the priority matrix. A total of 86 runs. After 20 times of operation, priority allocation of work area 1 is finished, and the result of priority allocation is 5; when the operation time is 26, priority allocation of work area 2 is finished, and the result of priority allocation is 7; after running to 34 times, priority allocation of work area 3 is finished, and the result of allocation priority is 8; when running to 86 times, priority allocation of job area 4 is finished, and the result of allocation priority is 29.

Therefore, the priority matrix of work area is as follows:

$$\begin{bmatrix} 5 & 7 & 8 & 29 \\ 5 & 7 & 8 & 29 \\ 5 & 7 & 8 & 29 \\ 5 & 7 & 8 & 29 \\ 5 & 7 & 8 & 29 \\ 5 & 7 & 8 & 29 \end{bmatrix}$$

In order to keep the balanced matching between the equipment group and the operation area, and the researchers use the priority to model the equipment group scheduling, two virtual operation areas are set, and the part that is not the square array is set to 0.

$$\begin{bmatrix} 5 & 7 & 8 & 29 & 0 & 0 \\ 5 & 7 & 8 & 29 & 0 & 0 \\ 5 & 7 & 8 & 29 & 0 & 0 \\ 5 & 7 & 8 & 29 & 0 & 0 \\ 5 & 7 & 8 & 29 & 0 & 0 \\ 5 & 7 & 8 & 29 & 0 & 0 \end{bmatrix}$$

The matrix is the final priority matrix generated by the 86 stage dynamic priority matrix in this paper. The priority of runway, taxiway, taxiway and apron are obtained as 5, 7, 8 and 29 respectively. The priority of ice and snow removal operation on

runway, taxiway, connecting way and apron is reduced successively, which is in line with the actual importance of the four pavements.

5 Conclusions

In this paper, firstly, the working area of airport pavement is divided into four areas: runway, taxiway, connecting way and apron. At the same time, the static priority model of operation area is established, and the weight confirmation method of information entropy is given. Through the dynamic priority algorithm, the dynamic priority of runway, taxiway, connecting track and apron is obtained As follows:

- (1) The remote sensing zoning method proposed in this paper can effectively partition the airport pavement de icing operation area, and make the airport pavement deicing operation area more information-based.
- (2) The initial priority obtained from the initial priority model of entropy weight is consistent with the reality of pavement importance.
- (3) By analogy with the differential positioning method in GPS measurement field, the proposed differential network time prediction method can be successfully applied to the dynamic priority algorithm.
- (4) Compared with the process scheduling algorithm of operating system, the dynamic priority assignment algorithm of job area generates 86 dynamic priority matrices in different stages in this paper. The priority changes dynamically with the job process, and the final priority generated is more reasonable than the initial priority.

Acknowledgements This research was funded by the key research and development program of Sichuan Province of China (nos. 2019YFG0050) and the science and technology cooperation research and development project of Sichuan Provincial Academy and University (nos. 2019YFSY0024).

References

1. Greg W, McCallum A (2018) Review of ice and snow runway pavements. *Int J Pavement Res Technol* 11(3):311–320
2. Sajed Sadati SM, Kristen C, Halil C, Alireza S, Kim S (2018) Energy and thermal performance evaluation of an automated snow and ice removal system at airports using numerical modeling and field measurements *Sustain Cities Soc* 43:238–250
3. Perrier N, Langevin A, Campbell JF (2008) The sector design and assignment problem for snow disposal operations. *European J Operat Res* 189(2): 508–525
4. Wilke S, Majumdar A, Ochieng WY (2014) Airport surface operations: A holistic framework for operations modeling and risk management *Safety Sci* 63:18–33
5. Labelle A, Langevin A, Campbell JF (2002) Sector design for snow removal and disposal in urban areas. *Socio-Econ Plann Sci* 36(3):183–202.

6. Angélica Salazar-Aguilar M, Langevin A Laporte G (2012) Synchronized arc routing for snow plowing operations. *Comput Operat Res* 39(7):1432–1440
7. Xu C, Shao Q (2018) Scheduling optimization of airport ground service support equipment under uncertain operation time. *Sci Technol Eng* 18(03):372–378
8. Bingliang N (2013) Research on data transmission technology for command and dispatching system of special vehicles in airport. College of Aviation Automation, Civil Aviation University of China, Tianjin
9. Hongjun H, Fang W (2017) Research on real-time scheduling of special vehicles in airport based on MAS. *Comput Appl Res* 34(09):2599–2604
10. Hongzhen Y (2013) Design of information system of airport ground deicing equipment based on GIS. College of aviation automation, Civil Aviation University of China, Tianjin
11. Ziyun R (2016) Research on cooperative scheduling of ground support vehicles for large airport flights. School of computer science and technology, Civil Aviation University of China, Tianjin
12. Xu S 2018 Research on collaborative scheduling modeling of large airport surface resources. Civil Aviation College, Nanjing University of Aeronautics and Astronautics, Nanjing
13. Fei C (2016) Research on airport deicing support technology under ice and snow conditions. *J Civil Aviation Univ China* 34(06):21–25
14. Li B, Yun B, Sudan, Xing Z, Zhu R (2012) Parallel machine scheduling method for aircraft deicing service resources. *Inf Control* 41(05):644–649
15. Xing Z, Tang G, Ren Z (2013) Noncooperative game of aircraft deicing process scheduling with multiple deicing pads. *Inf Control* 42(04):511–515
16. Yuanjun C (2016) Parameter prediction and operation monitoring of aircraft deicing equipment. School of Aeronautical Engineering, Civil Aviation University of China, Tianjin
17. Bin C, Liwen W (2019) Quantity optimization of aircraft ground deicing equipment based on flight dynamics. *J Syst Simulat* 31(04):757–764

Research on Object Detection and Location Using Semantic Segmentation



Huijie Zhu, Yan Cai, Yingchun Guo, Xiang Feng, and Wenya Zhang

Abstract Object detection and location generally require multiple sensors and multiple deep networks to achieve. For a specific application scenario, this paper proposes an algorithm based on semantic segmentation to display object detection and location at the same time. Due to the small size of the target in the image, especially the vertex of the target occupies very few pixels, the HRnet semantic segmentation algorithm with higher resolution is adopted. Due to the unbalanced class distribution and the small number of training samples, data augmentation and Lovasz Softmax loss function are used in a targeted manner. After experimental verification, in this data set, HRnet is superior to other algorithms. And by using data augmentation and Lovasz Softmax loss function, the accuracy and IoU have been greatly improved, especially the target vertex area with few samples.

Keywords Semantic segmentation · Lovasz softmax loss · Object detection · Object location

Object detection and location are research hot points in machine vision. They are widely used in fields such as unmanned driving, unmanned aerial vehicles and automation, and play a pivotal role in the development of artificial intelligence. The

H. Zhu · Y. Cai

Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China
e-mail: zhuhuijiehao@163.com

H. Zhu · Y. Guo

Science and Technology on Near-Surface Detection Laboratory, Wuxi 214035, China

X. Feng

The 28th Research Institution of China Electronics Technology Group Corporation, Nanjing 210007, China

e-mail: vonix@163.com

W. Zhang (✉)

School of Automation, Beijing Information Science and Technology University, Beijing 100101, India

e-mail: 13325177020@163.com

sensors used for object detection and location include cameras, lidars, millimeter wave radars, etc., among which camera has become the most widely used solution due to its low price and numerous vision algorithms. Object detection includes two tasks: object classification and image positioning. The positioning in object detection is mainly to determine the area of the object in the image, which is a two-dimensional solution problem. Object location is to calculate the distance between the object and the sensor. Since this information is not included in the image, it is generally necessary to use an algorithm to solve it indirectly. Object detection and location can finally obtain the three-dimensional information of the target.

With the application of deep learning in object detection and its accuracy rate is greatly ahead of traditional methods, it has become the mainstream of target detection. According to the different work steps, it is mainly divided into two categories: (1) Two-step object detection algorithm, this type of detection algorithm divides the detection problem into two stages. In the first stage, the candidate area is first generated, which contains the approximate location information of the object, and then the candidate area is classified and refined in the second stage. Typical representatives of this type of algorithm are Faster R-CNN [1], Cascade R-CNN [2] and TridentNet [3], etc. (2) One-step target detection algorithm. This type of detection algorithm does not require the area recommendation stage. It can directly generate the category probability and position coordinate value of the object through a network. Typical algorithms are YOLO, SSD and FCOS [4], etc. However, the rectangular frame of the object detection algorithm will contain a part of the background, and the object posture cannot be recognized. Therefore, it is difficult to accurately recognize the object, and the object location cannot be achieved.

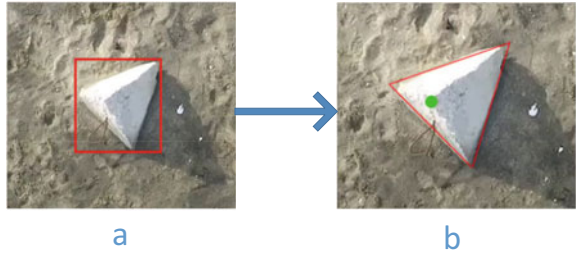
Semantic segmentation classifies each pixel, which can achieve more delicate and accurate recognition than object detection. Especially after the introduction of deep learning in recent years, the accuracy of segmentation has been greatly improved, and it has been widely used in unmanned driving, remote sensing images, industrial detection, etc. [5]. In this paper, by introducing semantic segmentation, the network is improved to enhance the recognition effect. At the same time, the size and shape of the object in the image are used to judge the distance between it and the camera, achieving detection and location. Furthermore, three-dimensional positioning is achieved by recognizing the top of the object.

1 Object Detection and Location Based on HRnet

1.1 Research Background

Triangular pyramids have multiple uses in road barriers, flood prevention and disaster relief, and military affairs. Its characteristics are similar to the external environment, and it lacks automated detection and location technology, which brings difficulties to maintenance and repair.

Fig. 1 Target detection and application requirements

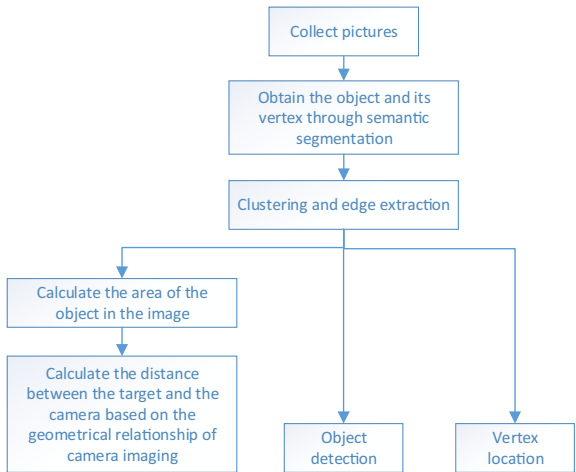


In practical applications, it is necessary to detect the position and shape of the triangular pyramid. Although the object detection can achieve the detection of the position and area, the accuracy of the detection method of the bounding box is insufficient, as shown in Fig. 1a. In addition, a triangular pyramid is an object with a certain height, and the exact position of its vertex needs to be detected in the application, as shown by the green dot mark in the Fig. 1b. If the lidar is applied, it will increase the load of the drone, and at the same time greatly increase the cost of the sensor. If the binocular camera is used, not only will it increase a lot of calculations, but the measurement distance is relatively short, and the accuracy is limited. Therefore, we develop a object detection and positioning scheme based on monocular camera.

1.2 Algorithm Framework

We use an algorithm based on semantic segmentation to achieve accurate detection and location of objects at the same time. The algorithm framework is shown in Fig. 2.

Fig. 2 Algorithm flowchart

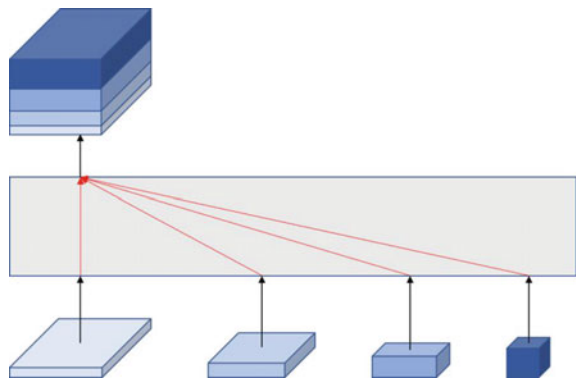


The collected pictures are input to the drone’s on-board processing unit or wirelessly transmitted to the ground station, and the semantic segmentation algorithm is used to segment the target area and vertex area, and then the accurate region of interest is obtained through the clustering algorithm and edge extraction. This region of interest realizes the detection of the target, and due to the target detection method of the bounding box, the detection accuracy is guaranteed. Since the triangular pyramid has a fixed size, its area will become smaller when the target is farther from the camera. Therefore, the distance between the camera and the triangular pyramid can be calculated according to the camera’s parameters and the region of interest, achieving the location of the object. The region of interest includes the position of the apex, which can be estimated by the distance between the pyramid and the camera and the height of the pyramid.

1.3 Introduction to HRnet

Semantic segmentation is one of the key issues of computer vision. Scene understanding is becoming more important, because more and more application scenarios in reality need to infer relevant knowledge or semantics from data. In view of the importance of semantic segmentation, many scholars have studied semantic segmentation algorithms using deep learning, such as Unet [6], Deeplab [7], PSPnet [8], ICnet [9], Fast-SCNN [10] and HRnet [11], etc. Different from other networks, HRnet always maintains high resolution during network transmission. The network repeatedly exchanges high-resolution low-scale information and low-resolution high-scale information, which allows the network to maintain a higher resolution while improving semantics. This feature makes it have outstanding performance in the fields of human pose estimation, semantic segmentation and object detection. Especially when the target is small, HRnet performance will be more prominent (Fig. 3).

Fig. 3 The framework of HRnet



This paper studies the detection and location of triangular pyramids. The vertices of the triangular pyramid are small in size in the image, and the detection will be particularly difficult when the target is far away. Therefore, we use the high resolution of HRnet to improve the effect of detection and location.

1.4 The Solution to Class Imbalance

Since the triangular pyramid occupies a relatively small proportion in the whole picture, the proportion of background pixels is much larger than the object. And the proportion of similar vertices of the pyramid is even smaller, which is generally less than 1% in the whole image. These will lead to class imbalance in the training process, making the recognition rate and IoU of the object and vertices low.

Although there have been methods such as focal [12] to automatically adjust the weight ratio of each category to solve the problem of class imbalance, the effect is still not ideal when the class distribution is severely unbalanced.

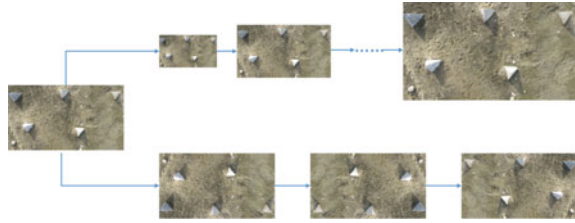
Lovasz loss is based on the convex Lovasz extension of submodular losses, which optimizes the mean IoU loss of the neural network [13]. A piecewise linear convex surrogate of the submodular set functions of Lovasz extension is a good optimized alternative to Jaccard loss. Lovasz loss can be divided into two types according to the number of categories of segmentation objects: lovasz hinge loss and lovasz softmax loss. Among them, lovasz hinge loss is suitable for binary classification problems, and lovasz softmax loss is suitable for multi-classification problems.

Lovasz softmax loss replaces the max-margin setting with logistic output, and uses the Softmax unit to map the score map output by the model to the probability distribution, similar to the traditional cross-entropy loss. First use the scoring function f to construct a pixel error vector, and then use it to construct a replacement function for loss. When evaluating the loss of all categories, the average value is used, so the Lovasz Softmax loss is used. Lovasz Softmax loss is generally used in combination with Softmax loss, and its ratio setting also has a certain impact on the network training effect. After a lot of experimental verification, in general, the network effect is better when Lovasz and Softmax account for 50%.

To a certain extent, Lovasz Softmax overcomes the problem that the back-propagation function adjusts less for the classes with small number, and improves the accuracy and IoU under class imbalance. We add this strategy to HRnet to further improve the training effect.

1.5 Data Augmentation

Due to the difficulty of collecting experimental samples and the high cost of labeling, the number of samples is small. Considering that the detection distance of the camera to the target is random, this will cause a large change in the size of the target in the

Fig. 4 Data augmentation

image, so we mainly use multi-scale expansion and change the image size by scaling. In addition, the viewing angle of the drone will also undergo various rotations, so on the basis of the scale change, the image are also mirrored and flipped up and down (Fig. 4).

2 Experiment Analysis

2.1 Experiment Introduction

We select pictures with representative class distribution from the collected samples for labeling, including 100 training samples and 20 test samples. Data augmentation is to randomly change the scale and angle of the samples during training. The resolution of pictures is 1280×720 , and the corresponding range of picture size after image enhancement is $640 \times 360 \sim 2560 \times 1440$. The standardization method is BN (batch normalization). The learning rate is 0.001. The pre-trained network adopts the pre-trained weight of the cityscapes dataset, using Adam as optimization algorithm, and the number of training times is 500. Magnitude settings in the three channels of HRnet are [18, 36], [18, 36, 72], [18, 36, 72, 144], and the weight of Lovasz Softmax loss and Softmax loss is [0.5, 0.5].

The computer configuration in the experiment: 64G RAM, NVIDIA 2080Ti GPU, 11G VRAM, Ubuntu 16.04 OS.

2.2 Experiment and Discussion

In order to verify the performance of each network in the experiment, first, only the Softmax loss function without data augmentation is used. The experimental results are shown in Table 1. Although PSPnet and Deeplab v3 are excellent semantic segmentation networks, the IoU of the triangular pyramid vertices is low due to the small target. Unet, ICnet, and Fast-scnn perform better in terms of computational efficiency. But ICnet and Fast-scnn retain less low-scale details, and their network fitting ability is limited, so their performances of accuracy and IoU are not good.

Table 1 Comparison of experimental results

network	IoU (triangular pyramid)	IoU (vertex)	kappa	acc (triangular pyramid)	acc (vertex)	Inference time (s)
Unet	0.7426	0.3128	0.8543	0.8573	0.4927	0.1368
Deeplab v3	0.7072	0.2686	0.8295	0.8434	0.4232	0.1570
PSPnet	0.7354	0.2484	0.8468	0.8600	0.4482	0.2063
ICnet	0.7242	0.2417	0.8403	0.8548	0.4755	0.1091
Fast-scnn	0.7418	0.2562	0.8524	0.8684	0.4707	0.1124
HRnet	0.7588	0.3197	0.8648	0.8601	0.5071	0.1808
HRnet ¹	0.7878	0.4094	0.8826	0.8883	0.5394	0.1970
HRnet ²	0.7621	0.3778	0.8671	0.8496	0.4912	0.1874
HRnet ¹⁺²	0.7949	0.4360	0.8864	0.8901	0.5694	0.1922

'1' represents using data augmentation, '2' represents using Lovasz Softmax loss

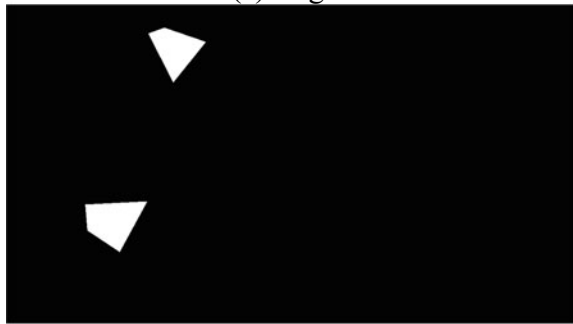
The U-shaped structure adopted by Unet retains some low-scale information, so it also has a good performance in target recognition, but it still has a big gap with HRnet. Therefore, low-scale semantic information is needed to improve the recognition accuracy of small objects, and high-scale semantic information is needed to distinguish the background from the object. HRnet continuously integrates low-scale and high-scale information in the network, so it performs better in this dataset. The result is shown in Fig. 5, where the red is the triangular pyramid and the green is the vertex of the pyramid. It can be seen that although the object is small, the shape of the triangular pyramid is still accurately recognized, and the vertex of the pyramid is accurately located. But the fly in the ointment is that a small stone near the triangular pyramid was also misdetected as the object, as shown in the yellow circle.

On the original HRnet network, the Lovasz Softmax loss function and the data augmentation method were also used for comparison. When using data augmentation, both IoU and accuracy have been significantly improved. This shows that the scale change simulates the change in the distance between the drone and the object, making the segmentation more accurate. When using the Lovasz Softmax loss function, the IoU of the triangular pyramid is improved slightly, but the vertices with smaller pixels have a significant improvement in IoU and accuracy. It is thus clear that Lovasz Softmax loss function has a significant improvement effect on the unbalanced class distribution. Finally, the simultaneous use of data augmentation and Lovasz Softmax loss for HRnet has significantly improved IoU, acc and kappa in each category. The prediction results are shown in Fig. 6, where the red is the triangular pyramid and the blue is the vertex of the pyramid. It can be seen that despite the messy cement blocks, stones and complex environment around, the object is still identified relatively accurately. Only when the object is very small as shown in Fig. 6a, the vertex missed occurs. Accurate detection and positioning can be achieved when the target is not very small, and the predicted triangular pyramid shape is consistent with the true shape.

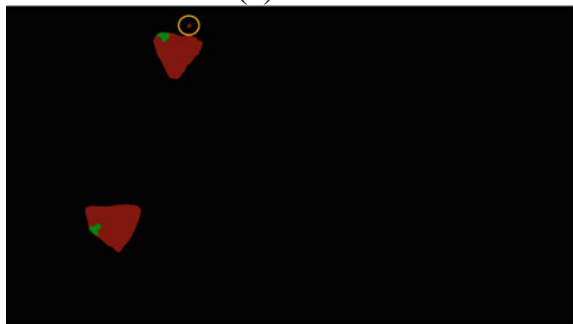
Fig. 5 Experimental results of HRnet



(a) origin



(b) label



(c) predict

3 Conclusion

This paper discovers that precise detection and location should be satisfied at the same time when triangular pyramid senses. Therefore, using semantic segmentation, a task that can efficiently and accurately achieve detection and location is developed, based on a monocular camera. HRnet combines high-scale semantics while maintaining low-scale, with high resolution. This experiment indicates that this method has advantages for the problem discussed in this paper. In view of the small amount of data and in order that the drone can be at any height, data augmentation algorithm

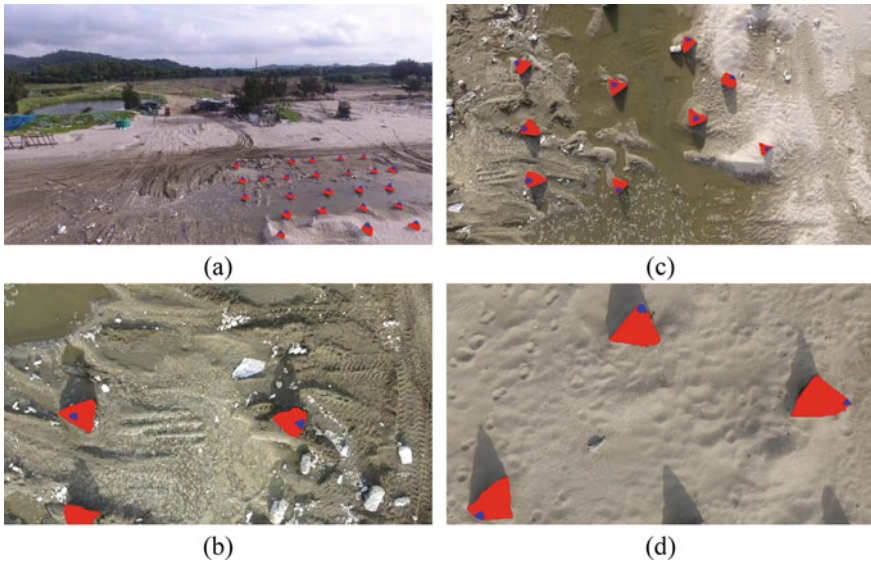


Fig. 6 Test results of the method in this article

is used. Furthermore, due to insufficient vertex samples, Lovasz Softmax loss is used to solve the problem of class imbalance. The experimental results show that the optimized algorithm has significantly improved the accuracy and IoU. This solution requires few sensors, and its accuracy meets actual needs, so it has high application value.

Acknowledgements This work was supported by Postdoctoral Science Foundation of China (Grant No. 2019 M661510), Science and Technology on Near-Surface Detection Laboratory (Grant No. 6142414180203, 614241419 0203), and the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 41704123).

References

1. Ren S, He K, Girshick R et al (2015) Faster r-cnn: Towards real-time object detection with region proposal networks: advances in neural information processing systems
2. Cai Z, Vasconcelos N (2018) Cascade r-cnn: Delving into high quality object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition
3. Li Y, Chen Y, Wang N et al (2019) Scale-aware trident networks for object detection. arXiv preprint [arXiv:1901.01892](https://arxiv.org/abs/1901.01892), 2019.
4. Tian Z, Shen C, Chen H et al (201) FCOS: fully convolutional one-stage object detection
5. L. D, M. Y, C. W et al (2018) Surround view cameras based instance segmentation method for autonomous vehicles. *Huazhong Keji Daxue Xuebao (Ziran Kexue Ban)/J Huazhong Univ Technol (Nat Edn)* 46(12):24–29

6. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer
7. Chen LC, Papandreou G, Kokkinos I et al (2018) DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans Pattern Anal Mach Intell* 40(4):834–848
8. Zhao H, Shi J, Qi X et al (2017) Pyramid scene parsing network. In: Proceedings of the IEEE conference on computer vision and pattern recognition
9. Zhao H, Qi X, Shen X et al (2018) Icnet for real-time semantic segmentation on high-resolution images. In: Proceedings of the European conference on computer vision (ECCV)
10. Poudel RP, Liwicki S, Cipolla R (2019) Fast-scnn: Fast semantic segmentation network[J]. arXiv preprint [arXiv:1902.04502](https://arxiv.org/abs/1902.04502)
11. Wang J, Sun K, Cheng T et al (2020) Deep high-resolution representation learning for visual recognition. *IEEE Trans Pattern Anal Mach Intell*
12. Lin TY, Goyal P, Girshick R et al (2017) Focal loss for dense object detection. *IEEE Trans Pattern Anal Mach Intell* (99):2999–3007
13. Berman M, Rannen Triki A, Blaschko MB (2018) The lovász-softmax loss: a tractable surrogate for the optimization of the intersection-over-union measure in neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition

General Standard Model for Safety Risk Assessment of Highway Bridge Construction



Yuqian Wang and Dianliang Xiao

Abstract With the high-level development of highway construction in China, it is very important to carry out standardized and universal risk assessment of bridge construction, and to find out the high-risk operations and risk event type distribution of the supervised bridges, by which is the basis for the related government departments to conduct focused risk supervision. This paper establishes general standard model for safety risk assessment of bridge construction in response to the departments' requirements of generality, standardized, convenient, easy to promote, and as objective as possible, as well as the limitations of current guideline "Safety Risk Assessment Guidelines for Highway Bridge and Tunnel Engineering Construction". This model is applicable to various risk events in different process for all types of bridges, including effective, reliable and measurable 14 first-level indexes and 82 second-level indexes based on 4M factors, which realized modularization and standardization of possibility assessment index system. After being verified by the application of Nansha bridge (the 2nd Humen Bridge) and many other bridges, the model is also used in the revision of "Safety Risk Assessment Guidelines for Highway Bridge and Tunnel Engineering Construction".

Keywords Highway bridge · Construction safety · Risk assessment · General standard model

1 Introduction

China's is still promoting her highway construction, when the main battlefield is gradually extending to large rivers and high mountains. The engineering geological conditions and construction environmental conditions of the project are more complex; the engineering technical difficulties are increasing; the spans of the bridges are also growing, so the projects' safety is facing severe challenges. In 2019, 67 accidents and 110 deaths were reported in the field of national highway and water

Y. Wang · D. Xiao (✉)
China Academy of Transportation Sciences, Beijing, China

transportation engineering construction, including 6 major accidents and 27 deaths, 1 serious accident and 12 deaths. Compared to 2018, the number of accidents increased by 25, 59.5%, and the number of deaths by 46, 71.9%. At the same time, the number of accidents in the first half of 2020 increased by 5.4% compared with the same period in 2019, and the number of fatalities fell by 4%. The construction safety of highway projects is still under high pressure, and it is urgent to use risk classification management to comprehensively improve the safety management level of enterprises and the safety supervision capabilities of government departments.

In China, bridge construction safety risk management is promoted from top to bottom. At the beginning, it was mainly the need of government supervision, which gradually led to the wave of enterprise spontaneous risk management. It is very important to carry out standardized and universal risk assessment of bridge construction, and to find out the high-risk operations and risk event type distribution of the supervised bridges, by which is the basis for the related government departments to conduct focused risk supervision. At the same time, the risk assessment model in the guides implies the government's acceptance of bridge construction safety risks and the summary of past accidents, which also plays an important guiding role in the spontaneous risk management of enterprises. Therefore, it is necessary to establish a general standard model for construction safety risk assessment in order to unify standards for the construction safety risk assessment of various bridge types in China, reduce the influence of the subjective factors of experts, and make the assessed risk level comparable across the country.

In order to establish a general standard model for highway bridge construction safety risk assessment applicable to various risk events in different process for all types of bridges in China, this article first investigates the common methods of bridge construction safety risk assessment at home and abroad, and analyzes the requirements for assessment method standardization, pointing out the limitations and improvement directions of the current guidelines for safety risk assessment of bridge construction. Then it expounds the research process of general standard model for safety risk assessment of highway bridge construction, describes the general standard model procedure and standard index database. Finally, application verification was carried out on Nansha Bridge (the 2nd Humen Bridge) and other large bridges.

2 Requirement Analysis

There have been many studies on bridge construction safety risk assessment methods at home and abroad. Among them, the International Tunnel Risk Management Guide issued by the International Tunnel Association (ITA) [1] was the most influential code in the field of civil engineering in China; many domestic engineering risk assessment codes have learned from this guide [2–5]. In recent years, the research trend is to use finite element calculation, fault tree, fuzzy comprehensive evaluation, cloud model, Monte Carlo and other methods to conduct risk assessment of specific risk events or a type of bridge [6–10]. The advantages of the above method are that the risk can

be obtained more accurately, and the construction company can be more effectively guided to adopt reasonable safety measures and appropriate organizational plans. The disadvantages are that the implementation is complex and requires a large amount of basic data, and it is difficult to use in other bridges which have different process and events. In short, they cannot meet the requirements of relevant government departments in focused risk supervision.

The needs of government risk supervision are mainly reflected in the following aspects:

- (1) The risk assessment method has to be easy to use, according to which the construction company with risk management experience can carry out the safety risk assessment of bridge construction without the help of a third party;
- (2) The risk assessment method is universal and standardized, so that it can be easily extended to various risk events in different process for all types of bridges;
- (3) Make the risk assessment conclusions as objective as possible, under the conditions of current expert experience and accident data.

3 Limitations of the Current Guideline

At present, the “Safety Risk Assessment Guidelines for Highway Bridge and Tunnel Engineering Construction” [2] is the main specification that guides China’s highway bridge engineering construction safety risk assessment. Compared with other qualitative risk assessment guidelines in civil engineering industry in the same period [3–5], it proposes an assessment method which combines quantitative index system method for possibility and qualitative expert consultation method for consequence, and reduces the impact of the subjective experience of experts. In the nearly ten years of application, the following shortcomings were also found:

- (1) Its possibility assessment index systems mainly target several common processes of numerous beam bridges, and is not suitable for other bridges which have different process and events.
- (2) Due to the uneven level of assessment companies, their Self-made risk assessment index systems are often missing the key risk sources, causing the accuracy of risk assessment to be questioned and mere formality.
- (3) According to the statistical analysis of production safety accidents in various industries in China, 80% of accidents are related to human factors, and 100% of accidents are related to management factors. However, the possibility assessment index system of human factors and construction management is relatively rough in the Guideline, not specific to each risk events.

Entrusted by the Ministry of Transport, China, the guide writing team improved the above shortcomings in the revision of the Guideline and proposed a general standard model for construction safety risk assessment of highway bridges.

4 General Standard Model

The general standard model of highway bridge construction safety risk assessment proposed in this paper has three main characteristics:

- (1) This model can be used for safety risk assessment of various risk events in different process for all types of bridges;
- (2) The construction method and index database of possibility assessment index systems are provided in a standardized way;
- (3) Through effective, reliable and measurable 14 first-level indexes and 82 second-level indexes based on 4M factors, the author summarized hundreds of experts' experiences in highway bridge construction safety risk assessment in recent ten years.

4.1 Research Process

In the aspect of risk identification, the construction processes and risk events of 156 highway bridge construction accidents in China in recent ten years were investigated. According to the number of accidents, 18 high-risk construction processes were identified, and the typical risk event distribution of each process was summarized, as shown in Table 1. According to expert opinions and investigation of Yangsigang Yangtze River Bridge (1700 m suspension bridge), Wufengshan Yangtze River Bridge (1092 m highway and railway suspension bridge), Nansha bridge (1688 and 1200 m suspension bridge), Shenzhen-Zhongshan Bridge Project (1620 m suspension bridge and 580 m cable-stayed bridge), 33 high-risk construction processes and risk events are supplemented, such as catwalk construction. According to the practices of safety risk identification for highway bridge construction in recent ten years, 416 safety risk sources of bridge construction were collected, which were classified as Human factors (268 in total), equipment and material factors (74 in total), environment factors (30 in total), management and method factors (41 in total).

In terms of risk assessment, the bridge engineering safety risk is modularized level by level according to typical risk events and risk sources. According to the terms of relevant Chinese Codes and the index set used in the safety risk assessment practice of highway bridge construction in the past ten years [11–14], the first-level indexes of each typical risk event are clarified, and the second-level assessment index database compiled, using theoretical analysis, literature and historical accident investigation, expert consultation and other methods.

Table 1 Distribution of high-risk construction processes and typical risk events

Construction process	Collapse/overtum	Lifting	Strike	Fall	Mechanical injury	Electric shock	Drowning	Vehicle injury	Poisoning choke	Explosion
Reinforcement work	○	○		○	○	○				○
Concrete casting			○	○				○		
Prestressed engineering			○		○					
Foundation pit construction	○	○	○	○						
Manual digging pile construction	○		○	○					○	
Bored pile construction		○	○							
Caisson construction	○	○		○			○	○		
Platform construction	○	○	○	○						
Pier construction	○	○	○	○						
The abutment construction	○	○	○	○						
Capping beam construction	○	○	○	○						
Support construction		○	○	○						

(continued)

Table 1 (continued)

Construction process	Collapse/overtum	Lifting	Strike	Fall	Mechanical injury	Electric shock	Drowning	Vehicle injury	Poisoning choke	Explosion
Template installation and removal	○	○	○	○	○					
Temporary facility installation and removal	○	○	○	○			○	○		
Cast-in-situ method of full supporting	○	○	○	○	○					
Bridge erection machine operation	○	○		○	○					
Cantilever cast-in-place operation	○	○	○	○						
Construction of guardrail and isolation pier				○	○	○				

Finally, a general standard model for safety risk assessment of highway bridge construction is established, including the standardized procedure of and the standardized assessment index database of highway bridge construction safety risk assessment.

4.2 Standardized Procedure

- (1) Risk identification. Identify key construction processes and risk events according to the identification table of construction processes and typical risk events; Identify the possible risk sources of each process and each risk event according to the construction safety risk factor analysis table.
- (2) Conduct consequence assessment. Use expert questionnaires to determine the level of consequences.
- (3) Possibility assessment. The first-level indexes are selected. According to the failure modes of risk events. Based on the relevant risk factors, project data, construction site and team investigation, appropriate second-level indexes are screened from the index database or supplementally created. Sort the importance of assessment indexes, determine the weight value and calculate the probability level.
- (4) Determine the risk level. According to the consequence level and probability level, the risk level can be obtained by risk matrix method.
- (5) Conduct risk control of risk level, classification, stratification and specialization, according to risk level, characteristics of risk events, risk level responsibility system and the construction safety risk factors table, so that the safety risk source of each post are determined and controlled (Fig. 1).

4.3 Index System Database

In order to establish the general standard model of construction safety risk assessment applicable to various bridge types, a standardized assessment index database was compiled, including effective, reliable and measurable 14 first-level indexes, 82 second-level indexes and their grading values based on 4M factors.

The first-level indexes of common risk events are shown in Table 2. The assessment indexes of human factors are shown in Table 3; the assessment indexes of facilities and equipment factors are shown in Table 4; the assessment indexes of environmental factors are shown in Table 5; and the assessment indexes of management factors are shown in Table 6.

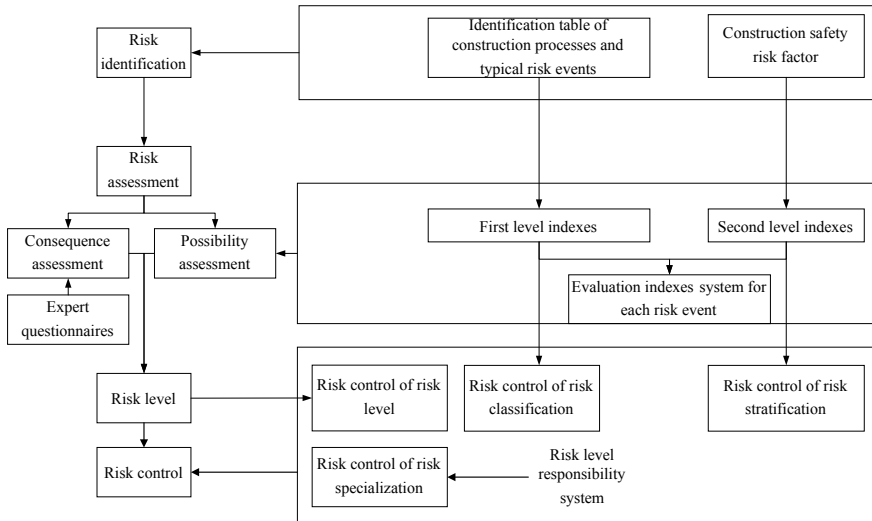


Fig. 1 Standardized procedure for highway bridge construction safety risk assessment

5 Case Study: Nansha Bridge

Using the general standard of highway bridge construction safety risk assessment model proposed in this paper, 28 construction processes and 170 risk sources of Nansha Bridge were extracted. Then the risk assessment were carried out for the collapse of foundation pit excavation, support collapse (piers of splay saddle, front anchorage chamber floor, Sidewall), support collapse (cross beams), tower column formwork collapse, catwalk collapse/fall during installation or removal, tower crane lifting injury, and steel box girder lifting injury. The universality and scientificity of the general standard model proposed are verified for various operating processes and risk events [15, 16].

6 Conclusion

This article mainly studies on the general standard model for safety risk assessment of highway bridge construction.

- (1) For focused safety risk supervision, relevant government departments need risk assessment method which is general, standardized, convenient, easy to promote, and as objective as possible.
- (2) This paper proposes a general standard model for safety risk assessment of highway Bridges construction to improve the limitations of the “Safety Risk Assessment Guidelines for Highway Bridge and Tunnel Engineering

Table 2 First-level indexes of common risk events

First-level indexes risk events	Safety awareness	Safety and emergency skills	Safety behavior or status	Infrastructure	Equipment	Workplace	Personal protection	Climatic environmental conditions	Geological and hydrological conditions	Surrounding environment	Management department	Management institution	Construction technology	Safety culture
Collapse (floor-standing structure)		○	○	○		○		○	○		○	○	○	
Collapse (floating structure)		○	○	○		○		○			○	○	○	
Collapse (equipment)		○	○	○	○			○	○		○	○		
Collapse (collision of vehicles or ships)			○	○				○	○		○	○	○	
Collapse (foundation pit)		○	○					○	○		○	○	○	
Fall	○	○	○			○	○	○				○		○
Lifting	○	○	○		○			○			○	○		
Strike	○	○	○		○		○	○		○	○	○		○
Mechanical injury	○	○	○			○		○				○		○
Electric shock	○	○	○			○	○	○				○		○
Drowning	○	○	○	○		○		○	○		○	○		
Vehicle injury	○	○	○	○		○		○			○	○		
Poisoning choke	○	○	○	○		○		○			○	○		
Explosion	○	○	○	○	○	○				○		○		

Table 3 Human factor index system

First-level indexes	Second-level indexes
Safety awareness	Level of education
	Pre-service training
	Regular training
	Technical clarification
Safety and emergency skills	Qualification of subcontractors
	Special operation certificates
	Subcontractor's related performance
	Emergency drills
Safety behavior or status	Safety officer staffing
	Health check
	Inspection by the owner or supervision company
	Inspection of construction company itself

Table 4 Index system of facilities and equipment factors

First-level indexes	Second-level indexes
Infrastructure	Foundation treatment
	Bridge alignment
	Electricity explosion-proof
	Anti-collision facilities
	Emergency facilities and equipment
Machinery and equipment	Advanced and applicable
	Span/arm length
	Lifting height
	Maximum lifting weight
	Track alignment
	Safety protection device
Workplace	Equipment management
	Operating space
	Ventilation
	Operation conditions
	Workplace toxicity
	Toxic substances
	Toxic workplace hours
	Frequency of fire source
Explosion hazard environment category	
Personal protection	Fire risk of combustible
	Personal protective equipment

Table 5 Index system of environmental factors

First-level indexes	Second-level indexes
Climatic and environmental conditions	Annual rainfall
	Extreme weather conditions
	Working height
	Wind conditions
	Foggy day
	Operating temperature
	Ambient mean humidity
Geological and hydrological conditions	Soil-rock conditions
	Geological conditions
	Groundwater
	Water depth
	Tidal range
	Trend
	Scour
Surrounding environment	Bridge site characteristics
	Topography
	Other hazards
	Typhoon shelter

Table 6 Index system of management factors

First-level indexes	Second-level indexes
Management department	Safety organization
	Department and post authority and responsibility
	Safety management goal planning
Construction technology	Advanced and applicable
	Temporary structure design
	Safety protection facilities
	Safety monitoring
	Level of automation
	Manual digging pile
	Foundation pit depth
	Support height
...	
Management institution	Institution of safety
	Assessment of institution
Safety culture	Safety culture

- Construction”, including mainly for several common processes of numerous beam bridges, incomplete assessment index system and rough consideration of human factors and construction management.
- (3) The general standard model for safety risk assessment highway bridge construction proposed in this paper is applicable to various risk events in different process for all types of bridges, including effective, reliable and measurable 14 first-level indexes and 82 second-level indexes based on 4M factors, which realized modularization and standardization of possibility assessment index system.
 - (4) The application was carried out on Nansha Bridge (the 2nd Humen Bridge), and the universality and scientificity of the general standard model proposed are verified. This model is also used in other large bridges such as Shenzhen-Zhongshan Bridge Project, and the revision of “Safety Risk Assessment Guidelines for Highway Bridge and Tunnel Engineering Construction”.

References

1. International Tunnel Association Working Group No. 2 (2004) Guidelines for tunneling risk managements. International Tunnel Association, Switzerland, pp 12–14
2. Engineering Quality Supervision Bureau of Ministry of Transport (2011) Safety risk assessment system and guidelines for highway bridge and tunnel engineering construction. China Communications Press, pp 231–232
3. CCCC Highway Planning and Design Institute Co., LTD (2010) Safety risk assessment system and guidelines for highway bridge and tunnel engineering design. China Communications Press, Beijing, pp 42–43
4. GB 50652-2011. Code for risk management of underground works in urban rail transit
5. Railway Construction (2007) No. 200, Provisional rules of risk assessment and management for railway tunnels
6. Cho T, Kim TS (2008) Probabilistic risk assessment for the construction phases of a bridge construction based on finite element analysis. *Finite Elem Anal Des* 44(6–7):383–400
7. Chai G, Ding S-S, Liu X-Y (2012) Risk assessment of bridge construction stage based on FCE. In: URKE 2012; international conference on uncertainty reasoning and knowledge engineering
8. Chen T-T, Wang C-H (2017) Fall risk assessment of bridge construction using Bayesian network transferring-from fault tree analysis. *J Civ Eng Manag* 23(2):273–282
9. Lu Z, Wei C, Liu M, Deng X (2019) Risk assessment method for cable system construction of long-span suspension bridge based on cloud model. *Adv Civ Eng* 9(5):5720637.1–5720637.9
10. Naderpour H, Kheyroddin A, Mortazavi S (2019) Risk assessment in bridge construction projects in Iran using Monte Carlo simulation technique. *Pract Period Struct Des Constr* 24(4):04019026.1–04019026.11
11. China Academy of Transportation Sciences (2019) Special risk assessment report on construction safety of Shenzhen-Zhongshan bridge project
12. The Research Institute of Highway, Ministry of Transport (2014) General research report on construction safety control and management of Maanshan Yangtze river highway bridge
13. China Academy of Transportation Sciences (2013) Research report on construction safety control methods of the Yellow River Bridge in Taohuayu, Henan
14. China Academy of Transportation Sciences (2019) Special risk assessment report on construction safety of the north section of Kaili Ring Expressway PPP project in Shandong province

15. China Academy of Transportation Sciences (2019) Safety risk identification, assessment and control manual for the suspension bridge construction of the 2nd Humen Bridge project
16. China Academy of Transportation Sciences (2019) Research report double prevention mechanism and technology of classified safety risk control and hidden risk screening and elimination in large highway bridge construction

Modeling Simulation and Stability Evaluation Analysis of Double-Trailer Combination



Yue Li, Hao Zhang, Hongwei Zhang, Chengqiang Zong, Chuanjin Ou,
and Xueli Zhang

Abstract In order to analyze the handling stability of double-trailer combination, the TruckSim simulation model was established, and the whole vehicle road test was carried out. The comparison between the road test results and the simulation results verifies the rationality and accuracy of the simulation model, indicating that the simulation model can reflect the characteristics of the double-trailer combination. By setting the steering wheel step input and the single line change test, the steering stability of the double-trailer combination was analyzed, concluding that Dolly is an important factor affecting the double-trailer combination. The simulation analysis results can provide the direction of improvement and technical guidance for the driving stability of double-trailer combinations.

Keywords Double-trailer combination · TruckSim simulation · Handling stability · Road tests

Y. Li (✉) · H. Zhang · H. Zhang · C. Zong · C. Ou · X. Zhang
Research Institute of Highway Ministry of Transport, Beijing, China
e-mail: li.yue@rioh.cn

Key Laboratory of Operation Safety Technology on Transport Vehicles of Ministry of Transport,
Beijing, China

H. Zhang
e-mail: zhang.hao@rioh.cn

H. Zhang
e-mail: hw.zhang@rioh.cn

C. Zong
e-mail: cq.zong@rioh.cn

C. Ou
e-mail: cj.ou@rioh.cn

X. Zhang
e-mail: xueli.zhang@rioh.cn

1 Introduction

Combination vehicles is a combination of one vehicle and one or more trailers [1]. The Ministry of Transport of the People's Republic of China, in May 2013, issued the "Guidance of the Ministry of Transport on Promoting the Healthy Development of the Logistics Industry" (Transport Planning and Development [2013] No. 349) [2], which for the first time proposed "to carry out applied technology research on double-trailer combination". Double-trailer combination are now widely used in Europe and other regions, China has issued the "General technical requirements for centre-axle trailer" (GB/T 37245-2018) and "General requirements for freight draw-bar trailers" (GB/T 17275-2019) and other relevant standards in terms of efficient transport vehicle models, and in 2018, the trial operation of centre-axle trailer combination was carried out [3], the introduction of the above regulations/policies, standards and demonstration operation work has laid a foundation for promoting the development of double-trailer combination in China.

For the study of double-trailer combination, domestic and foreign scholars have done related work. Sweatman and Tso [4] have simulated and tested B-double that it has a rollover characteristic similar to that of a six-axis semi-trailer combination, and the double saddle structure makes the B-double more stable.

Aoki et al. [5] established a theoretical analysis model of two combined models of type I (tractor + semi-trailer + semi-trailer) and type II (tractor + semi-trailer + full trailer), and studied its directional stability and controllability, concluding that under vibration, Type I is more stable than Type II, and the trajectory offset is slightly different from the steering sensitivity.

Ding and He [6] investigated the design and evaluation of the active steering system of a double-trailer combination based on driver-in-the-loop simulation, firstly the active steering control system of the trailer was designed, then the control system using LabVIEW and TruckSim was reconstructed, finally the reliability of the controller was verified by setting the 90° right-angle cornering condition and single-lane change condition.

Wang et al. [7, 8] established a four-degree-of-freedom dynamics model for the B-double train using the holistic and isolation methods, respectively, and found consistent results from the simulation comparison, simultaneously conducted a simulation analysis of the front wheel angle step input of the tractor, and obtained that the variation of parameters such as the transverse oscillation angle velocity response of each vehicle unit is consistent and more prone to instability under high speed conditions; they also proposed a new modeling method for the B-double train, and analyzed the relationship between the change of saddle position and the insufficient steering gradient of the three vehicle units, and proposed ideas for improving the structural design and operational stability of the vehicle.

In order to build software to imitate real research, domestic and foreign scholars have done related work. Li and Rakheja [9] proposed an active steering controller to prevent the train from collapsing when traveling on low-friction roads, where the upper level PID controller determines the corrective sway moment required for the

tractor and semi-trailer unit, and the lower level controller distributes the required sway moment for the tractor and semi-trailer, and the proposed controller is able to prevent the train from collapsing by performing open-loop sinusoidal input and J-shaped steering simulation analysis in TruckSim.

Douglas et al. [10] conducted a multi-body dynamics modeling study on commercial vehicles and used ADAMS to compare simulation and real-vehicle tests from the aspects of modeling, simulation, driver control, etc.

Yang et al. [11] established a non-linear simulation model for semi-trailer trains through TruckSim, built the TruckSim simulation model, and based on this, designed independent differential control controllers for tractors and semi-trailers.

Tang [12] investigated the lateral stability of a single sinusoidal input test semi-trailer car train by constructing a TruckSim whole car model.

Zhang [13] used TruckSim to construct a model of double-trailer trains (tractor + one trailer + tow platform + two trailer), and conducted a study on the active steering-in-the-loop control strategy through TruckSim-Simulink joint simulation control, and proved its effectiveness.

Huang [14] used TruckSim to model a mid-axle train from the body, steering system, powertrain and other systems, and investigated the steering stability of mid-axle trains by analyzing typical operating conditions such as steady-state rotation, serpentine and steering wheel step input.

Theoretical research on double-trailer trains in the above study, the semi-trailer trains with the help of software simulation, but the application of simulation software on double-trailer trains is less research, so, through software simulation on double-trailer trains modeling simulation, research and analysis of double-trailer trains operating stability and passability.

2 TruckSim Simulation Modeling

TruckSim is an integrated modeling, simulation and analysis software package developed by MSC, an automotive system simulation software development company. It is designed to simulate and analyze the dynamics of medium and heavy trucks, trailers, trains and buses, and has been adopted by many international automakers and parts suppliers. TruckSim can simulate the vehicle's response to the driver, road surface and aerodynamics. It is mainly used to simulate and analyze the stability, braking, smoothness, dynamics and economy of the vehicle, and is widely used in the development of modern vehicle control systems.

Based on TruckSim's user-friendly interface and parametric fast modeling features, this paper uses TruckSim to model the whole vehicle dynamics of A-double trains (tractor-trailer + Dolly + semi-trailer), which mainly includes the whole vehicle parameters, tire model, powertrain setup, axle system model, steering system model and load modeling 6 parts. The prototype parameters were provided by a company (Fig. 1).

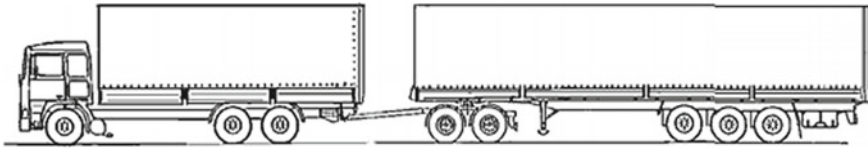


Fig. 1 Annotation of this figure

In the model solver of TruckSim software, there is no model solver ($S_{SS} + dSS + SSS$) corresponding to the models under study (3-axle tractor-trailer + 2-axle Dolly + 3-axle semi-trailer), the closest model is $S_{SS} + dS + S$, where S_{SS} indicates 3-axle tractor-trailer, dS indicates single-axle Dolly, and S indicates single-axle semi-trailer, so the axes (groups) of Dolly and semi-trailer need to be simplified to single axle.

2.1 Vehicle Body Modeling

The vehicle body is modeled in TruckSim software as the spring-loaded mass section is modeled, and the parameters for the three sections of the truck, Dolly and semi-trailer are shown in Table 1.

Table 1 Basic parameters for modelling the sprung mass of each vehicle unit

Parameter	Truck	Dolly	Semi-trailer
Length (mm)	10,333	4919	13,650
Width (mm)	2540	2000	2560
Height (mm)	3200	1270	2950
Longitudinal distance of the center of gravity from the coordinate origin (mm)	3560	3493	5596
Distance of the center of gravity from the origin of the coordinate (mm)	1350	1150	1890
Transverse distance of the center of gravity from the coordinate origin (mm)	0	0	0
Sprung mass (kg)	13,120	2730	11,315
Roll inertia ($kg\ m^2$)	9412.4	1894.3	8232.9
Pitch inertia ($kg\ m^2$)	75,697.2	8322.4	171,572.0
Yaw inertia ($kg\ m^2$)	75,697.2	8322.4	171,572.0
Product ($kg\ m^2$)	0	0	0
Product ($kg\ m^2$)	1626	0	0
Product ($kg\ m^2$)	0	0	6039

Table 2 Tyre parameters

Parameter	Numerical
Tire free diameter (mm)	1085
Effective rolling radius (mm)	527
Distance between two tires (mm)	2060/1870/1695/1840

2.2 Steering System Modeling

The nominal steering ratio selected for this vehicle modeling is 23:1. Other axle steering characteristics, elastic kinematics, asymmetric kinematics of the left and right wheels and steering wheel positioning parameters are default values.

2.3 Steering System Modeling

Tires are an important part of vehicle dynamics analysis, and their characteristics directly affect the dynamics, braking, handling stability, smoothness and safety of the vehicle. The number of tyres used in the prototype is 26, of which the front axle of the tractor-trailer and the semi-trailer Dolly double axle are single tyre, the rear double axle of the tractor-trailer and the triple axle of the semi-trailer are double tyre, the model number is 12R22.5 18PR, it can be seen from the standard “Size designation, dimensions, inflation pressure and load capacity for truck tyres” (GB/T 2977-2016), its single tyre load index is 152, the double tyre is 149, free diameter is 1085 mm.

The effective rolling radius of the tyre is calculated according to the recommended formula given by the European Tyre and Rim Technology Association (ETRTO).

$$r = F \times d / (2\pi)$$

where r is the rolling radius of the tyre.

F is the calculated constant, 3.05 for radial tyres and 2.99 for diagonal tyres.

d is the free diameter of the tyre (Table 2).

2.4 Axle System Modeling

As an unsprung mass, the axle system’s dimensions and mass, among other parameters, have a significant impact on the vehicle’s dynamics. The wheel core height and wheelbase are measured values, and the remaining parameters are the system default values, and the relevant parameters of the three vehicle units are shown in Tables 3, 4 and 5.

Table 3 Geometry and mass parameters of axle systems for truck

Parameter	Axle 1	Axle 2	Axle 3
Wheel core height (mm)	540	540	540
Wheelbase (mm)	2060	1870	1870
Axle center of gravity height (mm)	540	540	540
Unsprung mass (kg)	570	760	760

Table 4 Geometry and mass parameters of axle systems for Dolly

Parameter	Axle 1
Wheel core height (mm)	540
Wheelbase (mm)	1695
Axle center of gravity height (mm)	540
Unsprung mass (kg)	1280

Table 5 Geometry and mass parameters of axle systems for semi-trailer

Parameter	Axle 1
Wheel core height (mm)	540
Wheelbase (mm)	1840
Axle center of gravity height (mm)	540
Unsprung mass (kg)	1995

Table 6 Powertrain parameters

Parameter	Numerical
Maximum net engine power (kW)	360
Rated engine speed (rpm)	1900
Number of transmission gears	12

2.5 Powertrain Modeling

The prototype engine is a Foton Cummins ISGe5-490 diesel engine with a maximum power of 360 kW and a rated speed of 1900 rpm, with a maximum torque of 2300 Nm and a maximum torque of 1000–1100 rpm (Table 6).

2.6 Load Modeling

The load model selects the regular load, based on the constraints of combination vehicle compartment size, combination vehicle and unit mass, maximum design

Table 7 Regular load parameters of truck and semi-trailer

Parameter	Truck	Semi-trailer
Length (mm)	7820	13,600
Height (mm)	1000	1000
Width (mm)	2450	2450
Distance of mass center of load from front axle (mm)	3560	5596
Distance of load mass from ground (mm)	1600	1560
Distance of the load mass from the longitudinal symmetrical surface of the vehicle (mm)	0	0
Payload mass (kg)	9790	6680



Fig. 2 TruckSim simulation model

gross mass 49t and maximum design gross mass 25t for truck. The load model data for setting up tractor-trailers and semi-trailers is shown in Table 7.

The TruckSim simulation model established according to the above parameter settings is shown in Fig. 2.

3 Simulation Model Verification

In order to verify the validity of the established double-trailer train model and ensure that the simulation data and results are reliable and effective, the model is used to verify and analyze the model using the road test results of the prototype vehicle. By comparing and analyzing the simulation results with the test results, the validity of the simulation results is verified. Now, the vehicle designed and produced by an enterprise is used as the verification model, and the simulation model and the prototype vehicle are respectively subjected to steady-state rotation tests.

Under the same test conditions, the lateral acceleration and yaw rate of the traction truck and semi-trailer were selected as comparative evaluation indicators, and the simulation results were compared with the test results. The comparison results are shown in Figs. 3, 4, 5 and 6.

By comparison, it was found that the lateral acceleration of the tractor-trailer and the semi-trailer, the simulation results are consistent with the whole road test results curve change pattern, the curve fit is better, but the lateral acceleration value

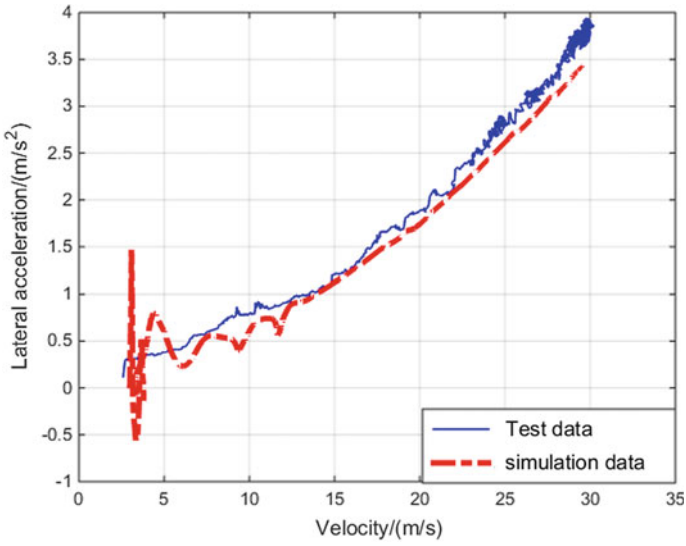


Fig. 3 Curve of lateral acceleration of truck

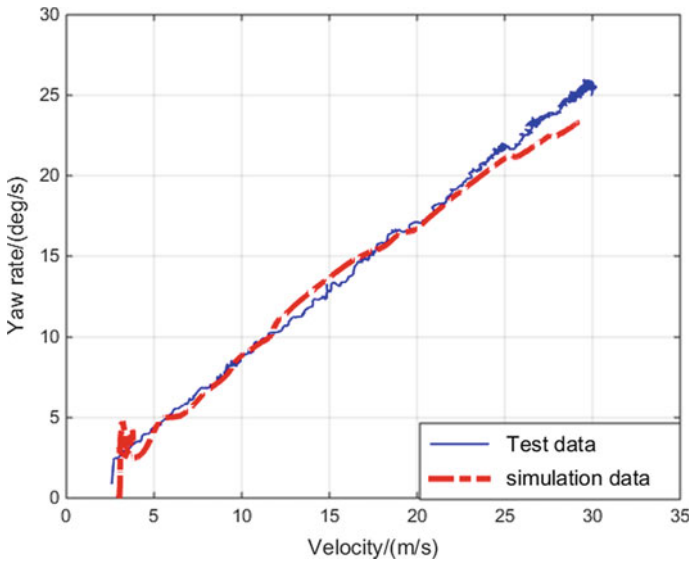


Fig. 4 Curve of yaw rate of truck

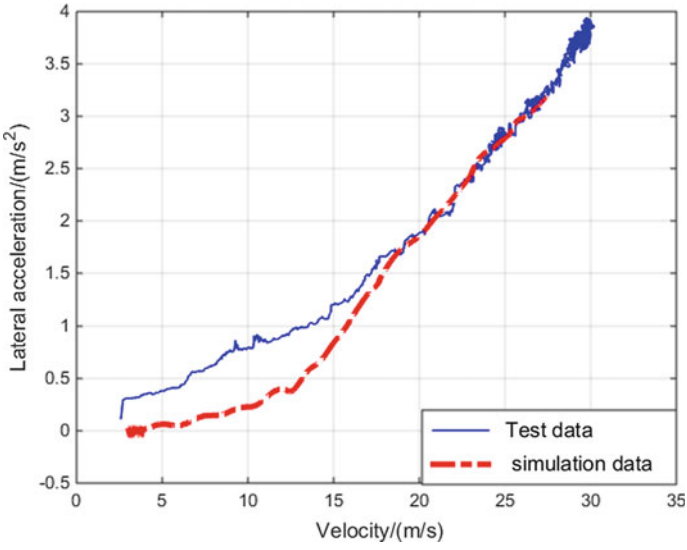


Fig. 5 Curve of lateral acceleration of semi-trailer

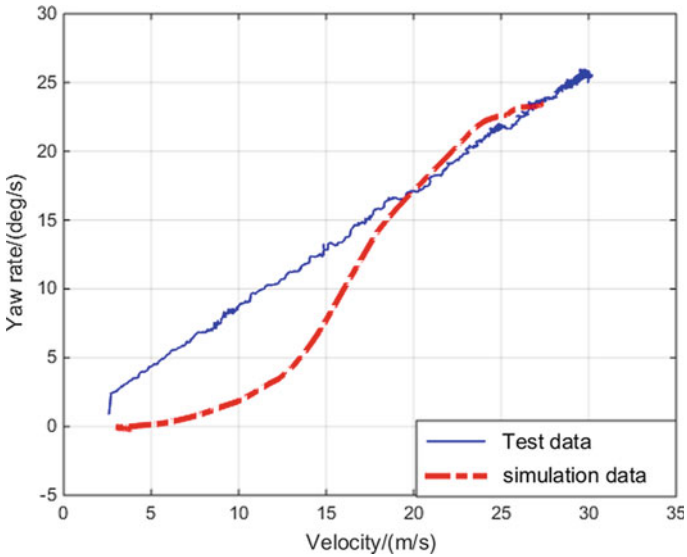


Fig. 6 Curve of yaw rate of semi-trailer

size has some error, probably because in the establishment of the simulation model, Dolly and the semi-trailer axle (group) is simplified to single-axis form. The above results illustrate that the simulation model can better represent the actual operation of the vehicle and reflect the good consistency, thus verifying the reasonableness and accuracy of the proposed model.

4 Simulation Analysis of Handling Stability

The handling stability of the vehicle refers to the ability of the vehicle to follow the direction given by the driver through the steering system and the steering wheel without excessive stress and fatigue, and to resist external interference and maintain stable driving [15].

The steering wheel step input test and single line change test were selected to analyze the stability of double-trailer combination.

4.1 Steering Wheel Step Input Test

The test was carried out in accordance with the test method given in “Controllability and stability test procedure for automotive” (GB/T 6323-2014) [16]. 80 km/h was selected and the steering wheel corner step input was 40° (as shown in Fig. 7) for the

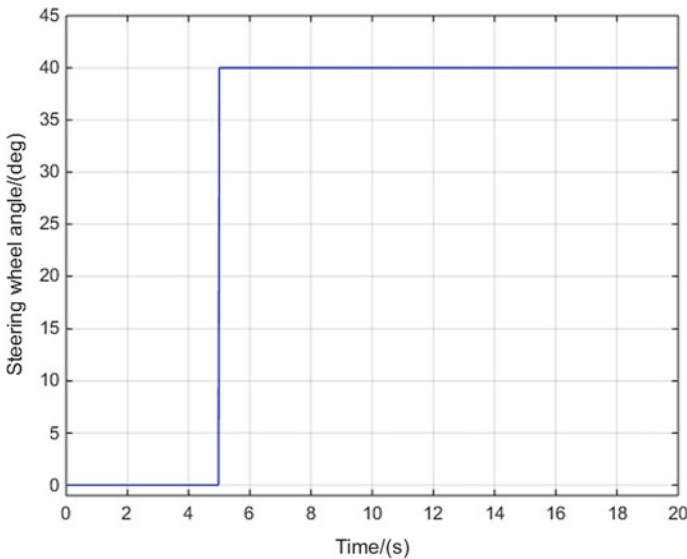


Fig. 7 Steering wheel angle step input

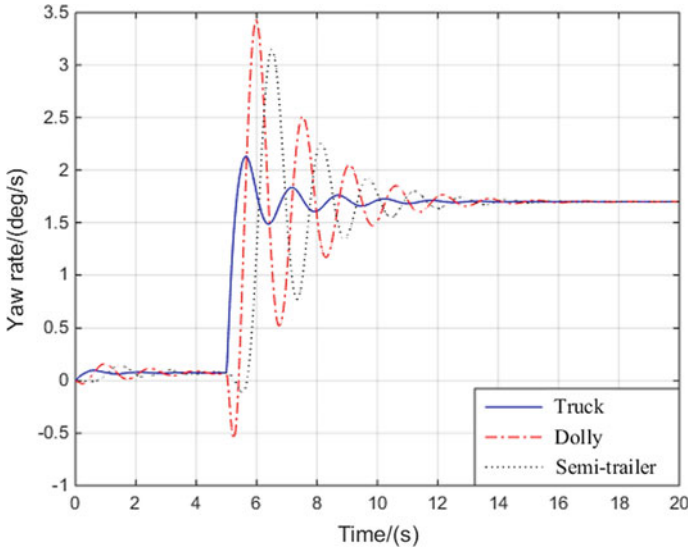


Fig. 8 Yaw rate response curve

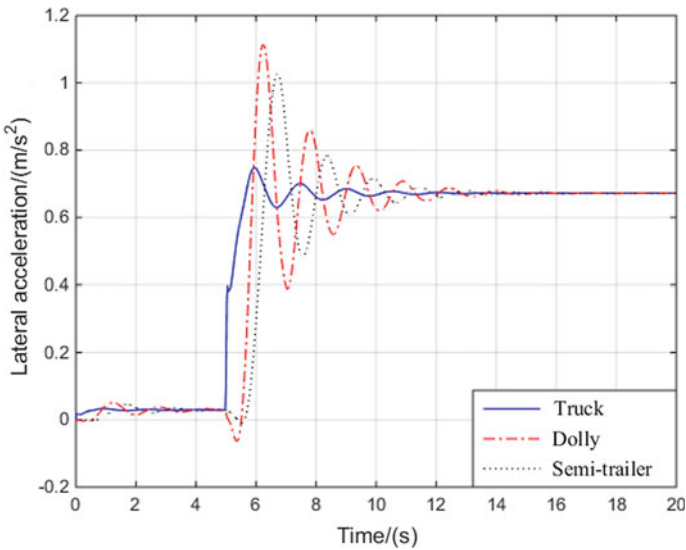


Fig. 9 Lateral acceleration response curve

test. After the completion of the train operation, each vehicle unit is gradually stabilized after a certain time of adjustment, the train adjustment process of the transverse oscillation angle speed, lateral acceleration, lateral inclination angle, lateral speed

and articulation angle response curve between vehicle units, as shown in Figs. 8, 9, 10, 11 and 12.

From Fig. 8, it can be seen that each vehicle unit's yaw rate reaches steady-state sway angle velocity value of 1.70 °/s after a certain period of time; from the moment

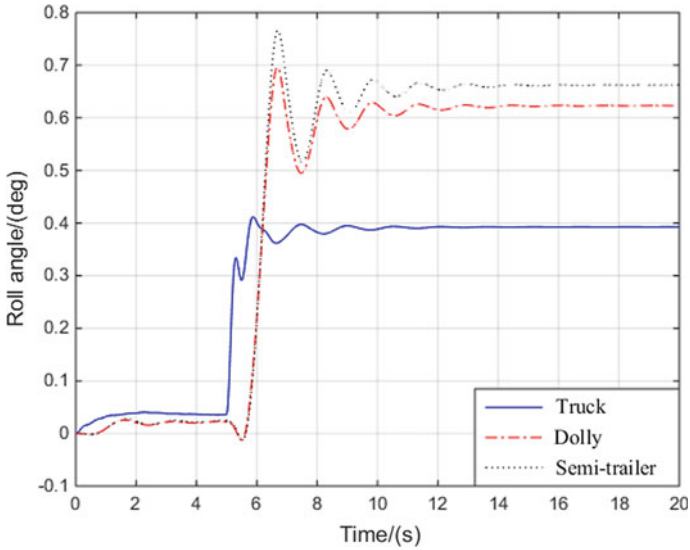


Fig. 10 Roll angle response curve

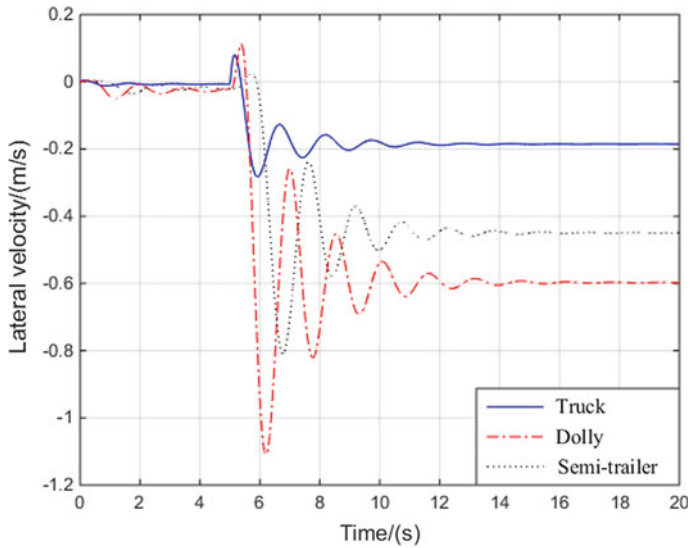


Fig. 11 Lateral velocity response curve

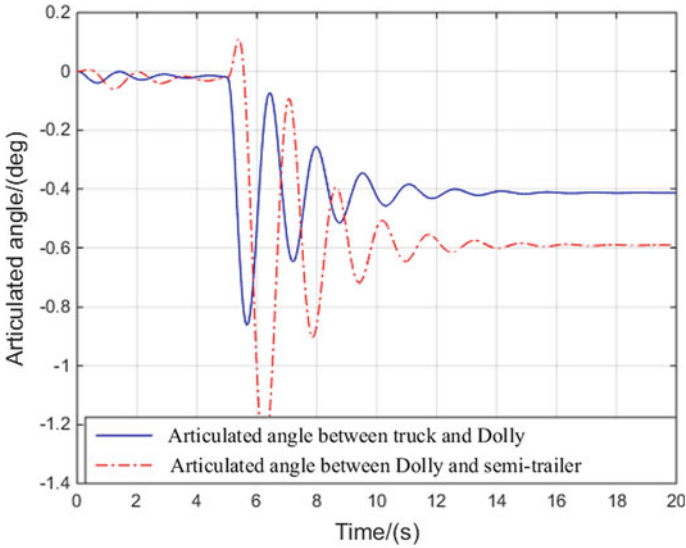


Fig. 12 Response curve of articulated angle between vehicle units

of the 5th second to apply the steering wheel angle step input, the moments when each vehicle unit reaches 1.70 °/s for the first time are 5.33 s, 5.61 s, 6.08 s, respectively, with time lag; the maximum value reached by each vehicle unit is 2.13, 3.43, 3.15 °/s, the amount of overshoot is 125.29, 201.76, 185.29%; the times when each vehicle unit’s sway angle velocity reaches steady-state value between 95 and 105% are: 8.05, 11.55, 11.35 s. From the peak of each unit, the amount of overshoot, and the time into steady state, it can be concluded that Dolly is the most active vehicle unit.

It can be seen from Fig. 9 that the maximum lateral acceleration of each vehicle unit is 0.75 m/s², 1.11 m/s², and 1.03 m/s² respectively, then the rearward amplification ratio of the lateral acceleration of the combination is 1.37, and there is a phenomenon of rear amplification. Dolly’s lateral acceleration value is the largest, and to reach steady-state time is the latest, indicating the existence of serpentine motion. It can be seen from Fig. 10 that when the roll angle of each unit reaches a steady state, the values are 0.39°, 0.62°, and 0.66° respectively. In the case of high-speed sharp turns, the rollover sequence is semi-trailer, Dolly, and traction truck.

It can be seen from Fig. 11, the maximum value of the absolute value of the lateral velocity of each vehicle unit is 0.28, 1.11, 0.81 m/s, the value of the steady state is 0.19, 0.60, 0.45 m/s, the maximum value of the lateral velocity of Dolly and the maximum value of the lateral velocity when reaching steady state is also the maximum, indicating the existence of serpentine motion. It can be seen from Fig. 12, the maximum value of the absolute value of the articulation angle between vehicle units is 0.86, 1.13°, the steady state value is 0.41, 0.59°, indicating the existence of vehicle rear amplification phenomenon.

4.2 Single Lane Change Test

The test was carried out in accordance with the test method given in “Road vehicles - Heavy commercial vehicle combinations and articulated buses - Lateral stability test methods” (GB/T 25979-2010) [17]. Select the test speed of 80 km/h, and apply a complete sine angle with a amplitude of 40° and a frequency of 0.25 Hz to the steering wheel, as shown in Fig. 13. The curves of the yaw rate, lateral acceleration, roll angle, lateral speed, and articulation angle between vehicle units with time are obtained, as shown in Figs. 14, 15, 16, 17 and 18.

It can be seen from Fig. 14 that the maximum values reached by each vehicle unit are 1.92, 2.79, and 2.75 °/s, and the time is 1.075, 1.350, 1.875 s, there is a time lag, and each vehicle unit Dolly’s last entered a steady state, indicating that Dolly’s yaw rate response is active.

It can be seen from Fig. 15 that the maximum lateral acceleration of each vehicle unit is 0.68, 0.96, 0.93 m/s², then the rearward amplification ratio of the lateral acceleration of the train is 1.36, and there is a phenomenon of rear amplification. And Dolly has the largest lateral acceleration value, and the latest time to reach the steady state, indicating the existence of serpentine motion. It can be seen from Fig. 16 that when the roll angle of each unit reaches a steady state, the values are 0.37°, 0.61°, and 0.68°, respectively. In the case of high-speed sharp turns, the rollover sequence is semi-trailer, Dolly, and traction truck.

It can be seen from Fig. 17 that the maximum absolute value of the lateral speed of each vehicle unit is 0.20, 0.69, 0.53 m/s, Dolly’s maximum lateral speed is the

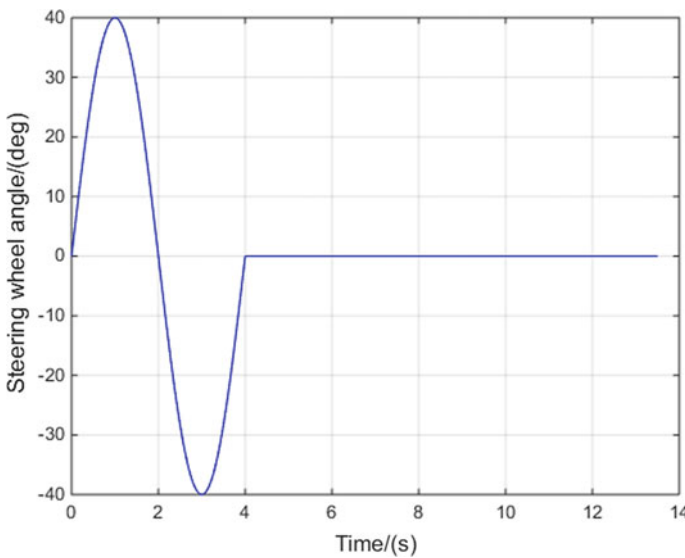


Fig. 13 Steering wheel angle change curve

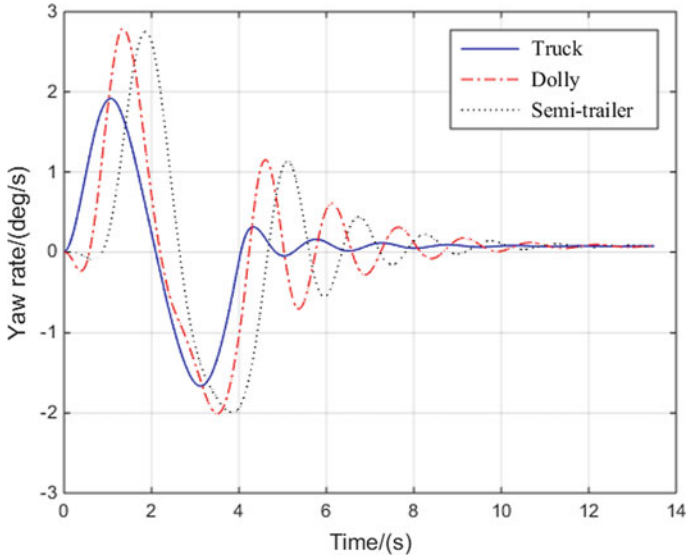


Fig. 14 Yaw rate response curve

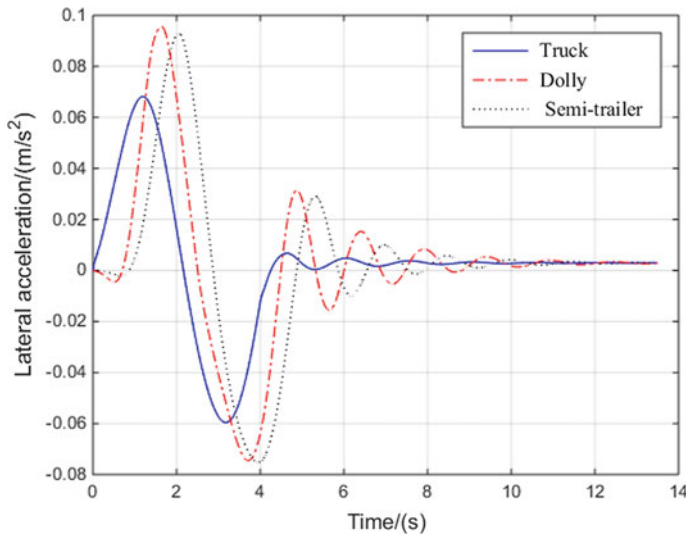


Fig. 15 Lateral acceleration response curve

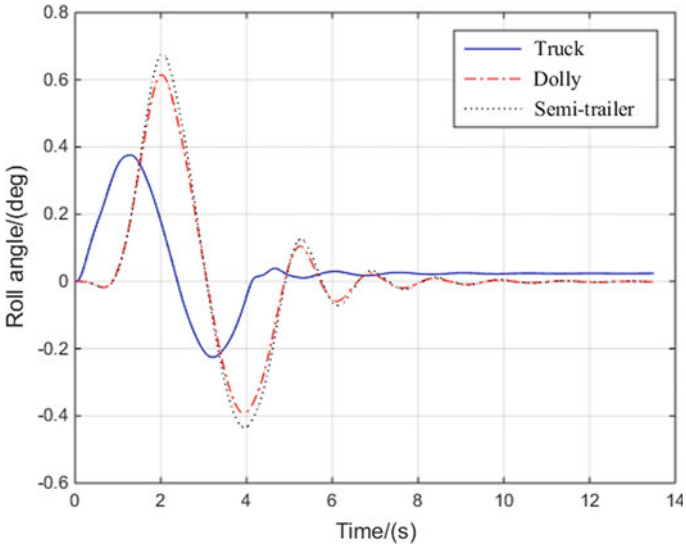


Fig. 16 Roll angle response curve

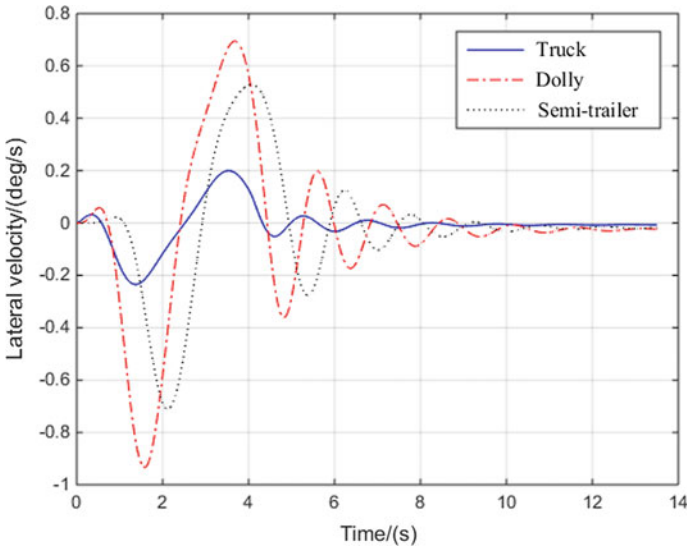


Fig. 17 Lateral velocity response curve

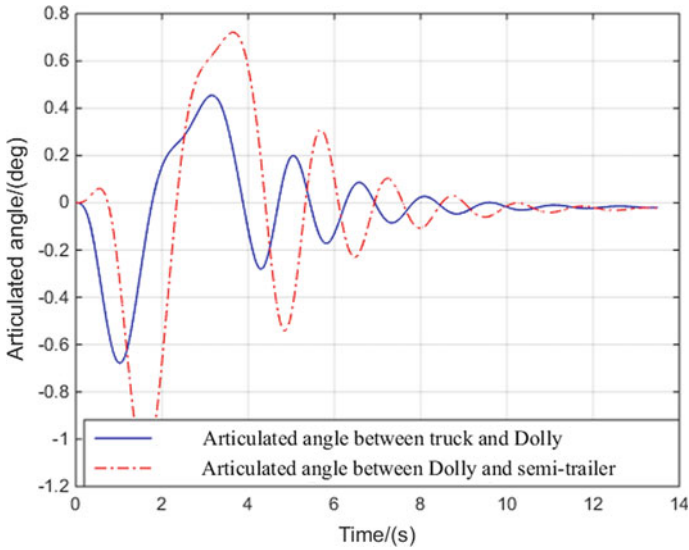


Fig. 18 Articulated angle response curve

largest and the steady state time is the latest, indicating the existence of Serpentine movement.

As can be seen from Fig. 18, the maximum value of the absolute value of the articulated angle between the vehicle units is 0.86 and 1.13°, and the steady state values are 0.41 and 0.59°, indicating that the vehicle has a rear magnification phenomenon.

5 Conclusion

In this paper, the dynamic simulation software TruckSim is used to build a model of a double-trailer train, and the rationality and accuracy of the built model are verified by comparing and analyzing with the road test data. The steering wheel angle step input test and single shift line test simulation conditions were set up to analyze and study the steering stability of the double-trailer train. It is concluded that Dolly is an important factor causing train instability, and it should be focused on when designing.

Acknowledgements This work is supported by transportation standard (quota) project No. 2019-10-073.

References

1. GB/T 3730.1 (2001) Motor vehicles and trailers—types—terms and definitions. Standards Press of China, Beijing
2. Ministry of Transport of the People's Republic of China (2013) Guidance of the Ministry of Transport on promoting the healthy development of the logistics industry. http://xxgk.mot.gov.cn/jigou/zhghs/201811/t20181115_3128872.html
3. General Office of the Ministry of Transport of the People's Republic of China (2018) Notice of the General Office of the Ministry of Transport on the development of the demonstration of modular center-axle trailer combination. http://xxgk.mot.gov.cn/jigou/yysfws/201806/t20180622_3036127.html
4. Sweatman P, Tso Y (1989) Dynamic stability of B-doubles. *Transp Plan Technol* 14(2):159–169
5. Aoki A, Marumo Y, Kageyama I (2011) Directional stability of multi-articulated vehicles with multiple axles. *J Mech Syst Transp Logist* 4(1):50–64
6. Ding X, He Y. Design and evaluation of active trailer steering systems of long combination vehicles using driver-software-in-the-loop simulations. SAE technical paper series. <https://doi.org/10.4271/2012-01-0934>
7. Wang GJ, Xu HG, Liu HF (2018) Comparative study on two methods for B-double vehicles. *J Hunan Univ (Nat Sci)* 45(02):53–61
8. Wang GJ, Xu HG, Liu HF (2019) Handling stability analysis of B-double. *J Zhejiang Univ (Eng Sci)* 53(02):299–306
9. Li B, Rakheja S. Jackknifing prevention of tractor-semitrailer combination using active braking control. SAE technical paper series. <https://doi.org/10.4271/2015-01-2746>
10. Douglas S et al (2002) Multi-body dynamic modeling of commercial vehicles. *Comput Control Eng* 13(1):11–15
11. Yang XJ, Kang N, Li XT (2013) Tractor-semitrailer stability control based on TruckSim-Simulink co-simulation. *J Highw Transp Res Dev* 30(03):141–148
12. Tang B (2015) Research on simulation analysis on lateral stability of tractor-semitrailer and comprehensive evaluation model. Research Institute of Highway M.O.T
13. Zhang YH (2017) Research on lateral instability mechanism analysis and in the loop control strategy of tractor and double-trailer combination. Jilin University
14. Huang CZ (2018) Research on handling stability and optimization of the centre axle trailer train. Research Institute of Highway M.O.T
15. Yu ZH (2014) Automobile theory. China Machine Press, Beijing, p 130
16. GB/T 6323-2014 (2014) Controllability and stability test procedure for automobile. Standards Press of China, Beijing
17. GB/T 25979-2010 (2010) Road vehicles—heavy commercial combinations and articulated buses—lateral stability methods. Standards Press of China, Beijing

An Improved Comprehensive Evaluation Method of Road Passenger Transportation Safety



Guoliang Dong, Fujia Liu, Haiying Xia, and Chaozhou Chen

Abstract Carrying out safety assessment before the departure of road passenger transport vehicles can find hidden dangers and prevent accidents. This article discusses the determination of the safety evaluation index of the transportation process, and the construction and improvement of the evaluation method. After analyzing the participating factors in the accident, the relevant factors affecting the safety of the transportation process are obtained. The main influencing factors were analyzed and combined with the principal component analysis method to analyze and filter the importance of the influencing factors to determine the safety evaluation index. Based on the comprehensive evaluation method, the road passenger transportation process safety evaluation method is constructed. Aiming at the problem of the inconsistency between the evaluation results of the traditional evaluation method and the actual input in the actual application, the reasons for the problems were analyzed and an improved comprehensive evaluation method was proposed. The improved evaluation method was applied to verify an accident instance, and the results showed that the improved evaluation method solved the problem of the deviation between output and input in the traditional evaluation method, and could objectively reflect the impact of various factors on the safety of the transportation process.

Keywords Road passenger transportation · Transportation process safety · Evaluation index · Evaluation method · Influencing factor · Improved evaluation method

1 Introduction

The safe operation of road passenger transportation vehicles is related to the life and property safety of the passengers. Evaluation of the safety status of the transportation process before the departure can eliminate hidden dangers and prevent accidents. A complete process from the starting station to the destination of the vehicle is called a

G. Dong (✉) · F. Liu · H. Xia · C. Chen
Research Institute of Highway Ministry of Transport, Beijing 100088, China
e-mail: gl.dong@rioh.cn

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_54

707

“transportation process”. This article discusses the factors affecting the safety of the transportation process, and proposes the construction and improvement of the evaluation method. After analyzing the participating factors in the accident and researching each factor, a relatively rough set of evaluation indicators is obtained. Applying principal component analysis to analyze the importance of influencing factors, the candidate index set can be screened and processed to determine the final evaluation index. The application of traditional evaluation methods often results in problems that the evaluation results do not match the actual input. Analyze the reasons leading to the above problems and propose an improved comprehensive evaluation method. The application of improved evaluation methods eliminates the contradiction between output and input and improves the scientific nature of evaluation. The improved evaluation method was applied to verify an accident instance and the results showed that the improved evaluation method eliminated the deviations caused by the traditional evaluation method and could more objectively reflect the impact of various factors on the safety of the transportation process.

2 Analysis of Safety Factors

In order to obtain more scientific evaluation indicators, analyzed and studied the accident participation factors, accident statistics, accident investigation reports and other related data of passenger transportation vehicles. Analyzed and researched the factors involved in the accident, and extracted the main influencing factors.

2.1 Analysis of Accident Participation Factors

Researching the statistical data of the main causes of traffic accidents of operating vehicles, it can be found that the main causes of motor vehicle fault accidents are: driver’s fault, mechanical failure, road factors, etc., covering three major factors such as drivers, vehicles, and roads. The proportion of traffic accidents caused by driver’s fault is over 85%.

Figure 1 is a statistical data chart of the main causes of accidents caused by motor vehicle violations. It can be seen that the factors that account for the larger number of accidents are: speeding, illegal driving, illegal traffic signals, retrograde, overtaking, illegal meeting, driving without a license, illegal occupation of the road, and fatigue driving.

According to the accident statistics, the main participating factors of the operating vehicles are sorted out, as shown in Table 1. The items in the table cover the aspects of driver, vehicle and management.

Analyzing the statistics of commercial passenger car accident cases, it can be found that the direct causes of commercial vehicle accidents mainly include improper operation, speeding, crossing the line, overtaking, mechanical failure, puncture,

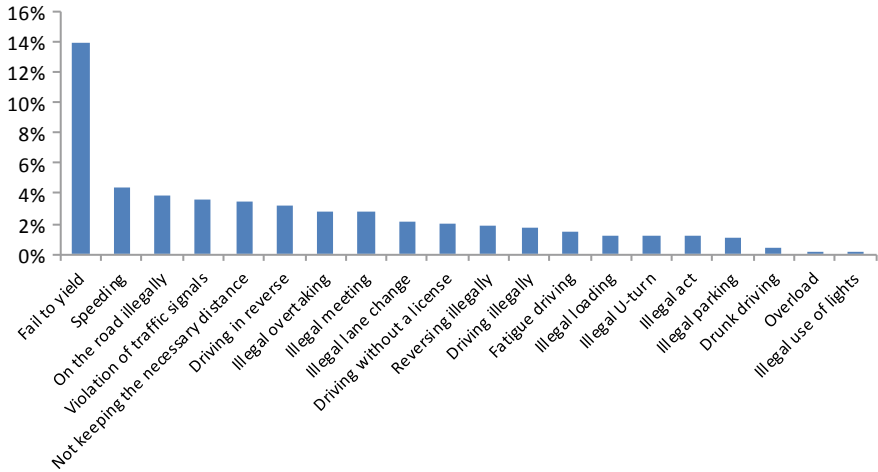


Fig. 1 Main cause of vehicle accidents

Table 1 Participation factors of the operating vehicle

No.	Cause of accident	No.	Cause of accident
1	Speeding	13	Illegal parking
2	Drunk driving	14	Driving illegally
3	Driving in reverse	15	Illegal loading
4	Fatigue driving	16	Overload
5	Illegal lane change	17	Violation of traffic signals
6	Illegal overtaking	18	Fail to yield
7	Reversing illegally	19	Driving without a license
8	Illegal U-turn	20	Illegal use of lights
9	Illegal meeting	21	Not keeping the necessary distance
10	Illegal traction	22	Improper braking
11	Illegal act	23	Improper steering
12	On the road illegally	24	Improper throttle control

fatigue driving, overloading, red light running, etc. The number of accidents caused by drivers exceeds 50%, so driver factors are the main factor in accidents.

Relevant studies have shown that the key influencing factors of automobile transportation safety are mainly the overall quality of the driver, vehicle maintenance, etc. [1]. Among the causes of accidents, driver factors have the greatest impact, followed by environmental factors and vehicle factors [2]. Studies have shown that the driver’s subjective factors play an important role in the safe operation of vehicles, and about 90% of road traffic accidents are caused by human factors [3]. Drunk driving, fatigue driving, speeding and disobeying the regulations are the four driver factors that have

the greatest impact on road traffic accidents [4]. Therefore, the influencing factors can be classified and studied from the aspects of driver, vehicle, road environment, management, etc.

2.2 Driver Factor

Driver factors include many indicators such as age, driving experience, personality, etc. The following is a study of the main influencing factors on driving safety.

Age. Figure 2 shows the statistical distribution of the “age” indicator in the accident case statistics. It can be seen that that drivers aged 40–50 have the most accidents, followed by the 30–35, 35–40, and 50–55 age groups.

Related studies have found that drivers under 30 have the highest traffic accident rate, and those around 50 have the lowest accident rate [5]. Drivers aged 6–45 have more accidents [2]. Younger drivers are at higher risk [6], and those under 25 are at greater risk [7]. The age of accidents is mainly concentrated between 28 and 41 years old [8].

Based on the above research, the following conclusions are drawn: Drivers under the age of 25 have the highest driving safety risk. Between 25 and 40 years old, as the age increases, the risk gradually decreases. Drivers between the ages of 40 and 49 have the least safety risks. Over the age of 50, as the age increases, the risk gradually increases.

Driving experience. Analyzing the statistical data of the “driving experience” indicator in accident cases, we can find that the number of accidents with driving experience of 15–20 years accounted for the vast majority. Based on the relevant research results [2, 4, 8–11], the following conclusions are obtained. (1) The relationship between driving experience and safety is not linear. (2) Drivers who have been

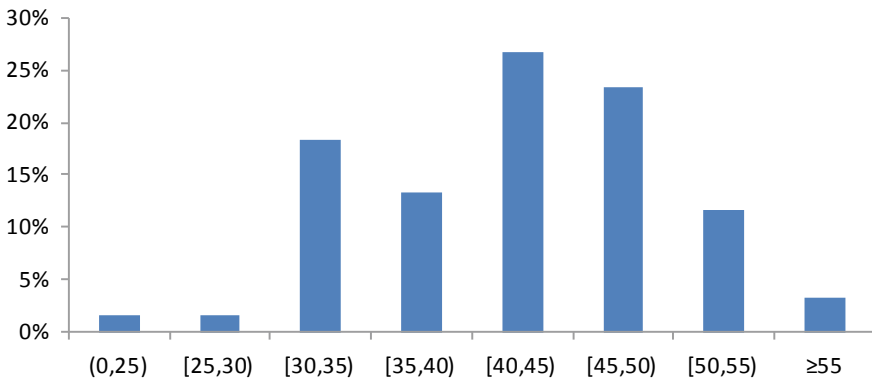


Fig. 2 Distribution of accidents by driver's age

driving for less than 5 years have the greatest safety risk. For the driving for 5–10 years, with the increase of driving experience, safety risks gradually increase. If the driving experience is 10–12 years, the risk gradually decreases as the driving experience increases. For those who have been driving for more than 12 years, the risk decreases with age.

Driving skills. Driving skills are the external performance of the driver's safety ability. Safety awareness affects risk perception ability and driving skills, and driving skills are affected by risk perception ability [12]. Driving skills are an objective reflection of the driver's skills and are related to the process safety of the vehicle [13]. The impact of this indicator on safety should be considered in the evaluation.

Safety consciousness. Analyzing the statistical data of accident cases, it can be found that the driver factor is the main cause of the accident. The driver's subjective factors play an important role in the safe operation of the vehicle. Drivers' weak safety awareness leads to serious violations, which is one of the main reasons for road traffic accidents [1, 3, 14]. Safety awareness, risk perception and driving skills are the main causes of traffic accidents caused by drivers, and safety awareness affects risk perception and driving skills [12]. Therefore, it is necessary to strengthen the cultivation of driver's safety awareness and increase its weight in the evaluation model.

Safety education. Analyzing the statistics of accident cases, it can be found that many accidents are caused by the driver's weak safety awareness. Drivers' weak safety awareness is a major hidden danger that leads to road traffic accidents, and the development of safety awareness is a subtle result [15]. Risk education methods can significantly reduce the rate of dangerous traffic incidents. The driver's risk perception ability after training has been greatly improved compared to before. The monthly education will reduce the risk of dangerous traffic incidents first and then rise [16–18].

2.3 Vehicle Parts and Age

Statistics show that the proportion of accidents caused by vehicle technical conditions is less than 30% [2]. Analyzing the causes of accidents caused by vehicle component failures, it is found that many factors such as poor vehicle conditions, inadequate vehicle maintenance, and poor safety awareness, have combined to cause the accident. Mechanical failures mostly occur in parts such as brakes, steering, and tires.

The age of the vehicle is an inherent indicator of the vehicle. As the age of the vehicle grows, the technical status of the vehicle is gradually decreasing. Analyzing the statistics of vehicle age indicators in accident cases, it can be found that vehicles with a vehicle age of 4–6 years have the largest proportion of accidents.

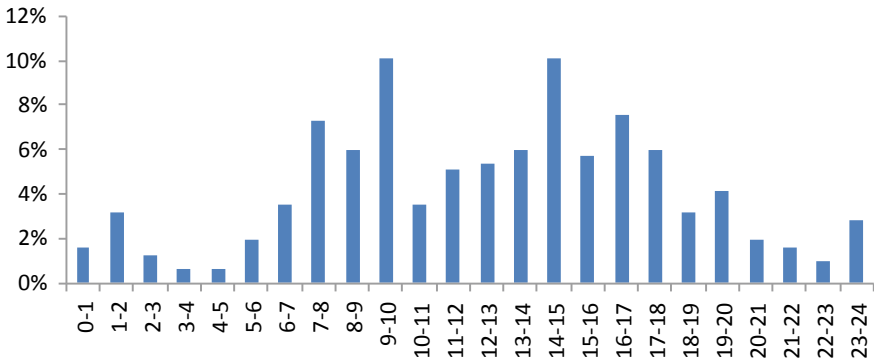


Fig. 3 The distribution of accidents by time period

2.4 Road Environmental Factors

Accident prone time. Figure 3 shows the accident statistics data of accident cases by time period. It can be seen that accidents are prone to occur at 7–10 and 11–18, with the highest peaks at 9–10 and 14–15. The driving speed at night fluctuates more than during the day, and the accident rate at 14–18 o'clock in the day is much higher than other time periods [2, 19]. Therefore, 1–3 o'clock, 11–13 o'clock, and 17–19 o'clock can be designated as accident-prone time periods.

Road conditions and weather reasons. It can be seen from the statistical data of the “road conditions” indicators in the accident cases that in the unfavorable road conditions, the number of accidents on water pavement accounts for the vast majority, followed by the conditions of icy and snowy pavements. Weather conditions such as rain, snow, and fog will adversely affect the driver’s line of sight and road conditions, and affect the safe driving of vehicles. Therefore, consideration should be given to the impact of unfavorable conditions such as water, snow, icing, rain, snow, and fog on driving safety.

2.5 Principal Component Analysis Method for Influencing Factor Analysis

Among the factors affecting traffic accidents, there may be correlations between factors. More independent variables will complicate the evaluation model. The principal component analysis method can be used to extract key variable information and find out the main influencing factors under the classification of drivers, vehicles, roads, and management. The main conclusions are as follows:

The main influencing factors related to the driver: the driver’s safety awareness, speeding, failing to maintain a safe speed, failing to observe the road conditions effectively, improper operation in danger or improper handling after an accident.

The main influencing factors related to the vehicle: braking system, vehicle maintenance.

Influencing factors related to the road and environment: steep slopes/long downhills, sharp bends, slippery roads, rain, snow and fog.

The main influencing factors related to management: business methods, satellite positioning monitoring, company safety awareness, company supervision of drivers and vehicles.

The main influencing factors obtained by the principal component analysis method are verified by the regression method. The comparison chart (part) of the regression curve and the original value curve is shown in Fig. 4. It can be seen that the fitting results of the impact of various factors on the accident are basically consistent with the original data. Therefore, the accident influencing factors obtained by the principal component analysis method are effective.

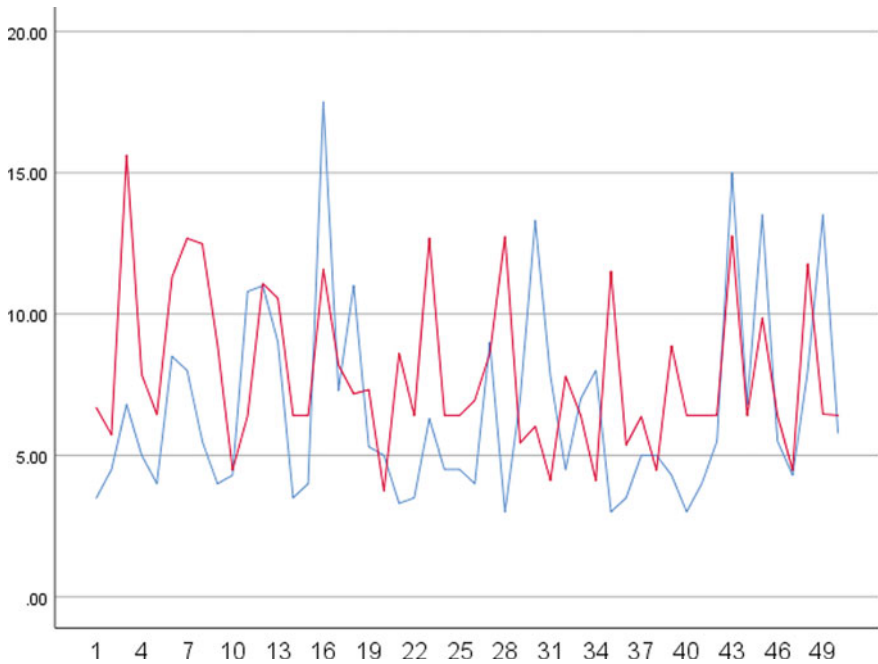


Fig. 4 Comparison of the regression and the original

3 Evaluation Method

3.1 Choice of Evaluation Method

Evaluation methods such as expert scoring method, analytic hierarchy process and comprehensive evaluation method are applied in various fields [20–24]. The comprehensive evaluation method has the advantages of simple calculation and reasonable evaluation results. Therefore, the comprehensive evaluation method is selected for the safety evaluation of the transportation process.

3.2 Membership Function

The membership function represents the degree of membership of each evaluation factor to each evaluation level. The methods to determine the membership function mainly include expert survey method, assignment method, fuzzy statistics method, etc. According to the different characteristics of the evaluation index, different methods are selected to determine the membership function in the safety evaluation of transportation process. For qualitative indicators, the expert survey method is used to score the importance of the qualitative indicators by experts to construct fuzzy subsets of membership. For indicators such as driving style, maintenance conditions, weather conditions, and road conditions, expert methods is used to construct fuzzy subsets of membership. For quantifiable indicators, the assignment method is used to determine the membership function. Quantitative indicators such as driving age, age, vehicle age, and vehicle mileage can set corresponding scoring rules, and then use the membership function to convert the evaluation score into a subset of membership. The membership function is based on the classic fuzzy distribution, and the value of each parameter in the function is calculated and determined according to the preset curve shape. Commonly used fuzzy distributions include rectangular distribution, trapezoidal distribution, normal distribution, Cauchy distribution and so on. Based on the Cauchy distribution in the safety evaluation of transportation process, the membership function curve of each evaluation level is determined according to the construction principle of membership degree, as shown in Fig. 5. The construction principles of membership functions include: the curve is convex, the same point does not take the maximum value in the two membership functions at the same time, and the same point should not belong to more than two regions at the same time. When constructing the membership function, the error between the reverse calculation value and the input value of the membership subset should also be verified to ensure that the error is kept within a reasonable range (for example, the error value is within 3). Table 2 lists the deviation between the deblurred calculated value and the original input value at some sampling points. The absolute maximum deviation is 1.3, which is within a reasonable range.

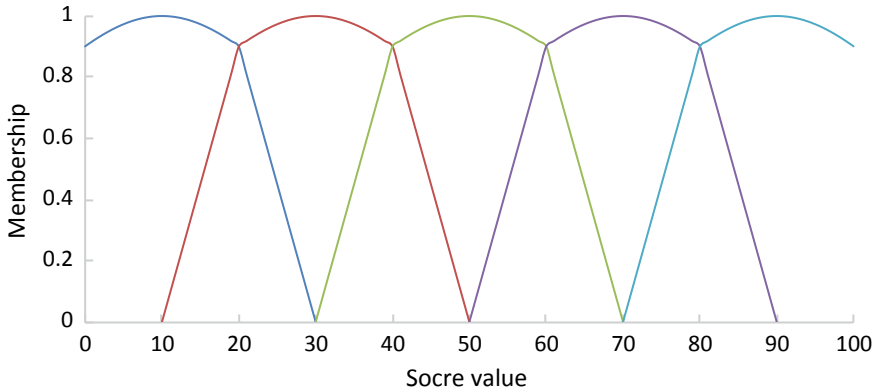


Fig. 5 Membership function graph

Table 2 Original input value (OIV), membership degree subset, deblurred calculated value (DCV), deviation between DCV and OIV

No.	OIV, x	Uv1(x)	Uv2(x)	Uv3(x)	Uv4(x)	Uv5(x)	DCV, x1	x1 - x
1	2	0.200	0.000	0.000	0.000	0.000	2.0	0.0
2	8	0.800	0.000	0.000	0.000	0.000	8.0	0.0
3	15	0.684	0.360	0.000	0.000	0.000	16.3	1.3
4	20	0.500	0.500	0.000	0.000	0.000	20.0	0.0
5	30	0.000	1.000	0.000	0.000	0.000	30.0	0.0
6	80	0.000	0.000	0.000	0.500	0.500	80.0	0.0
7	85	0.000	0.000	0.000	0.316	0.684	83.7	-1.3
8	95	0.000	0.000	0.000	0.000	1.056	95.0	0.0
9	100	0.000	0.000	0.000	0.000	1.111	100.0	0.0

3.3 Traditional Evaluation Model

The safety evaluation index of the transportation process is divided into 3 layers. Vectors A, R, and B respectively represent the weight value, evaluation matrix and membership degree of each layer. The traditional evaluation model generally requires that the value of each index in the sub-layer should be mutually exclusive, and the sum of the weight value of each index should be 1. In the evaluation calculation, from the third layer to the first layer, the comprehensive evaluation calculation is performed layer by layer to the upper layer until the evaluation value of the target layer is obtained. The mathematical model of each layer of indicators is as follows:

Mathematical model of the third layer:

$$\left\{ \begin{array}{l} A_{11} = (a_{111}, a_{112}, a_{113}) \\ B_{11} = A_{11} \cdot R_{11} \\ A_{12} = (a_{121}, a_{122}, a_{123}, a_{124}, a_{125}, a_{126}) \\ B_{12} = A_{12} \cdot R_{12} \\ A_{13} = (a_{131}, a_{132}, a_{113}, a_{134}, a_{135}, a_{136}, a_{137}) \\ B_{13} = A_{13} \cdot R_{13} \end{array} \right. \quad (1)$$

$$\left\{ \begin{array}{l} A_{21} = (a_{211}, a_{212}) \\ B_{21} = A_{21} \cdot R_{21} \\ A_{22} = (a_{221}, a_{222}, a_{223}, a_{224}, a_{225}) \\ B_{22} = A_{22} \cdot R_{22} \\ A_{23} = (a_{231}, a_{232}, a_{233}) \\ B_{23} = A_{23} \cdot R_{23} \end{array} \right. \quad (2)$$

$$\left\{ \begin{array}{l} A_{31} = (a_{311}, a_{312}, a_{313}, a_{314}, a_{315}, a_{316}, a_{317}, a_{318}) \\ B_{31} = A_{31} \cdot R_{31} \end{array} \right. \quad (3)$$

Mathematical model of the second layer:

$$\left\{ \begin{array}{l} A_1 = (a_{11}, a_{12}, a_{13}) \\ R_1 = (B_{11}, B_{12}, B_{13})^T \\ B_1 = A_1 \cdot R_1 \end{array} \right. \quad (4)$$

$$\left\{ \begin{array}{l} A_2 = (a_{21}, a_{22}, a_{23}) \\ R_2 = (B_{21}, B_{22}, B_{23})^T \\ B_2 = A_2 \cdot R_2 \end{array} \right. \quad (5)$$

$$\left\{ \begin{array}{l} A_3 = (a_{31}, a_{32}, a_{33}, a_{34}) \\ R_3 = (B_{31}, R_{32}, R_{33}, R_{34})^T \\ B_3 = A_3 \cdot R_3 \end{array} \right. \quad (6)$$

$$\left\{ \begin{array}{l} A_4 = (a_{41}, a_{42}, a_{43}, a_{44}, a_{45}) \\ R_4 = (R_{41}, R_{42}, R_{43}, R_{44}, R_{45})^T \\ B_4 = A_4 \cdot R_4 \end{array} \right. \quad (7)$$

Mathematical model of the first layer:

$$\left\{ \begin{array}{l} A = (a_1, a_2, a_3, a_4) \\ R = (B_1, B_2, B_3, B_4)^T \\ B = A \cdot R \end{array} \right. \quad (8)$$

Evaluation result k:

$$k = B \cdot C = (b_1, b_2, \dots, b_n) \cdot \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix} = \sum_{i=1}^m b_j \cdot c_j \tag{9}$$

According to the section to which the evaluation result k belongs in the evaluation level, the safety evaluation level of this transportation process is determined.

3.4 Improved Evaluation Model

When the traditional evaluation model is used to evaluate the safety of the actual transportation process, there are often problems that the evaluation results are seriously inconsistent with the actual situation. For example, the brake system of the vehicle is abnormal, the tire is damaged, and the road is downhill. But the evaluation conclusion may be “no obvious risk”, which is seriously inconsistent with the actual situation. After analyzing the calculation process of the evaluation model, it is found that the main reasons are as follows.

- (1) In the traditional evaluation model, the sum of the weight values under the sub-layer of the evaluation index is required to be 1. A larger weight value more objectively reflects the importance of the indicator. But when the layer-by-layer stacking calculation is performed, the degree of influence of high-weight indicators will be greatly weakened. Especially when multiple main indicators are concentrated under the same sub-layer, the requirement that the sum of weights is 1 seriously weakens the importance of key indicators. In view of this situation, in the improved model, the sum of weight indicators under the sub-layer is no longer limited to 1.
- (2) The index value of the conventional evaluation model is generally a single option, which cannot meet the actual requirements in the safety evaluation of the transportation process. According to the actual situation, the value should be set to multiple options, and necessary methods should be taken when multiple merging to ensure the rationality of the merged result.

The following takes the index of “key vehicle components” as an example to illustrate the consequences and improvement methods caused by the above two problems. The key components of the vehicle are the third-layer indicators, and their values are shown in Table 2. The abnormality of the key components of the vehicle may be caused by the problems of multiple components in the steering, braking, and tires at the same time, or it may be a problem of a single component. Therefore, the single-option method is unreasonable. After analyzing the key factors affecting the accident, it can be found that these key components have a great impact on vehicle safety. Therefore, higher weight values should be set for these 3 indicators, for example, the weights are all set to 0.9. When three components have problems at

Table 3 Key vehicle components

Indicator name	Value
Key vehicle components	Steering system abnormal
	Abnormal braking system
	Damaged or severely worn tires
	Normal

the same time, if the traditional model calculation method is used, due to the limitation of the sum of weights, the weight value of the key components of the combined vehicle may still be 0.9, which greatly weakens the impact of multiple component failures on the vehicle. In the improved evaluation model, assuming that the sum of the weights is set to 2, the combined value of the key components is 2. Compared with traditional evaluation methods, the importance of evaluation indicators in the improved evaluation method has been greatly improved, which is more in line with the actual situation (Table 3).

3.5 Instance Verification

In order to verify the actual effect of the evaluation method, an accident case is used for verification.

Basic situation. In an accident, a bus crossed the central guardrail and collided with a small bus. Many people on the bus were thrown out of the bus, resulting in 8 deaths, 1 serious injury. The bus took full responsibility for the accident. After investigation, the bus was operated for 5.5 years, with severe tire wear, lack of a safety hammer, and some safety belts damaged. The bus driver was 33 years old, 14 years of driving experience, and 13 illegal records. The accident section was a straight section with a smooth road. When the accident happened, it was light rain. The road was wet. There was no water, fog, and no obstructions that would affect the sight. The driver did not remind the members to wear seat belts, and the passenger transport company did not monitor the vehicles well. The following uses an improved evaluation method to calculate the safety evaluation of the accident.

Third-layer evaluation calculation. For the driver’s indicator, the importance coefficient of each factor is:

$$A_{11} = (0.355 \ 0.383 \ 0.263)$$

The driver’s physical fitness is calculated as “average” and the evaluation score is 75. The driving experience is 14 years and the evaluation score is 72. The age is 33 years old and the evaluation score is 74. Then:

$$R_{11} = \begin{bmatrix} 0.000 & 0.000 & 1.000 & 0.000 & 0.000 \\ 0.000 & 0.359 & 0.641 & 0.000 & 0.000 \\ 0.000 & 0.153 & 0.847 & 0.000 & 0.000 \end{bmatrix}$$

$$B_{11} = A_{11} \cdot R_{11} = (0.000 \ 0.177 \ 0.823 \ 0.000 \ 0.000)$$

The same method is used to calculate the evaluation results of other indicators.

$$B_{12} = A_{12} \cdot R_{12} = (0.000 \ 0.058 \ 0.877 \ 0.065 \ 0.000)$$

$$B_{13} = (0.000 \ 0.000 \ 0.000 \ 0.000 \ 1.221)$$

$$B_{21} = A_{21} \cdot R_{21} = (0.000 \ 0.028 \ 0.814 \ 0.158 \ 0.000)$$

$$B_{22} = (0.000 \ 0.000 \ 0.000 \ 1.297 \ 0.703)$$

$$B_{23} = A_{23} \cdot R_{23} = (0.000 \ 0.000 \ 0.000 \ 0.360 \ 0.640)$$

$$B_{31} = (0.000 \ 0.000 \ 1.000 \ 0.000 \ 0.000)$$

Second-level evaluation calculation. For the driver's indicator, the calculation steps are as follows.

$$A_1 = (0.1980.3350.468)$$

$$R_1 = \begin{bmatrix} 0.000 & 0.177 & 0.823 & 0.000 & 0.000 \\ 0.000 & 0.058 & 0.877 & 0.065 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 1.221 \end{bmatrix}$$

$$B_1 = A_1 \cdot R_1 = (0.000 \ 0.049 \ 0.414 \ 0.020 \ 0.517)$$

Calculate the evaluation value of the vehicle index in the same way.

$$B_2 = A_2 \cdot R_2 = (0.000 \ 0.006 \ 0.173 \ 0.657 \ 0.527)$$

Evaluation value of road environment index is:

$$B_3 = A_3 \cdot R_3 = (0.000 \ 0.000 \ 0.512 \ 0.253 \ 0.235)$$

Evaluation value of management factor index is:

$$B_4 = A_4 \cdot R_4 = (0.000 \ 0.000 \ 0.000 \ 0.383 \ 0.617)$$

First-layer evaluation. The calculation process of this transportation safety evaluation is as follows.

$$A = (0.400 \ 0.267 \ 0.100 \ 0.233)$$

$$R = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix} = \begin{bmatrix} 0.000 & 0.049 & 0.414 & 0.020 & 0.517 \\ 0.000 & 0.006 & 0.173 & 0.657 & 0.527 \\ 0.000 & 0.000 & 0.512 & 0.253 & 0.235 \\ 0.000 & 0.000 & 0.000 & 0.383 & 0.617 \end{bmatrix}$$

$$B = A \cdot R = (0.000 \ 0.019 \ 0.240 \ 0.272 \ 0.469)$$

Determine the security level. The evaluation score of the target layer is:

$$k = B \cdot C = \sum_{i=1}^{mn} b_j \cdot c_j = 95.4$$

Therefore, the risk evaluation of this transportation process before the accident was 95.4 points and the risk level was “very high”. The evaluation conclusion is consistent with the actual situation.

4 Conclusion

After analyzing and researching historical passenger accident cases and accident statistics, the influencing factors of accidents are analyzed and a rough evaluation factor set is constructed. Combined with relevant historical research results, the key influencing factors related to the accident are analyzed in focus. Using big data processing methods, the importance of influencing factors is studied, and the safety evaluation index of the transportation process is established. Aiming at the problem that the traditional evaluation method may cause the evaluation conclusion to be inconsistent with the actual situation, the cause of the problem is analyzed, and an improved evaluation method is proposed. In order to verify the effectiveness of the improved evaluation method, a safety evaluation calculation is performed on an actual accident, and the evaluation conclusion is “very high” risk, which is consistent with the actual situation.

Acknowledgements This paper was funded by the Special Fund Project for Science and Technology Innovation of Research Institute of Highway Ministry of Transport (2019-C507).

References

1. Wu B, Sun Q, Guo Y et al (2018) Analysis of key influencing factors of automobile transportation safety management. *Logist Eng Manag* 040(001):158–160
2. Zhu S (2016) Research on the traffic safety risk of operating buses based on the accident tree analysis. Chang'an University
3. Gao L (2015) Research on the evaluation method of driver's safety awareness simulation driving. Kunming University of Science and Technology
4. Zhang L, Liu T, Pan F et al (2014) Analysis of the influence of driver factors on road traffic accident indicators. *Chin Saf Sci J* 24(5):79–84
5. Lie X (2016) Analysis of the influence of driver's age on driving safety. *Energy Conserv Environ Prot Transp* 6:15–17
6. Shi J, Peng D, Gao Y (2018) Research on improper driving behavior of motor vehicle drivers with 12 points deducted. *Traffic Inf Saf* 36(06):30–37
7. Yuan L, Yuan H, Wang W (2013) Analysis of the influence of driver's age on traffic risk perception ability. *Soc Psychol Sci* 6:91–94
8. Shi N (2018) Analysis of the frequency distribution of driver accidents and its influencing factors. Beijing Jiaotong University
9. Liu X (2016) Research on the correlation between drivers' driving experience and traffic accidents. Chang'an University
10. Ding Y, Chang Y, Gao Y (2014) Analysis of characteristics of traffic accidents of drivers of different driving ages. *Traffic Inf Saf* 32(5):198–202
11. Liu W, Zhou Z, Niu Q et al (2015) Comparative study on safety management of drivers with low driving age at home and abroad. *Road Traffic Saf* 5:36–42
12. Tao J (2018) In-depth analysis of drivers of commercial vehicle traffic accidents. *Shanxi Transp Sci Technol* 253(04):155–158
13. Yang P (2013) Research on the safety evaluation method of operating drivers. Chang'an University
14. Zhang Q (2010) Research on driver safety awareness evaluation. Chang'an University
15. Li X, Zeng C, Cai F et al (2012) Research on driver's safety awareness experiential training device. *Energy Conserv Environ Prot Transp* 8(3):51–56
16. Wang X, Luo Y, Yu R et al (2016) Traffic safety education for operating drivers based on risky driving behavior analysis. *Automob Saf* 1:62–66
17. Fan D, Cao K (2010) The impact of driver's risk awareness on traffic safety. *Chin Saf Sci J* 20(11):30
18. Luo M, Wang X, Yu R et al (2016) Research on the effect of behavioral safety education of operating drivers. *Chin Saf Sci J* 26(9):124–128
19. Wen J, Yang X, Wu C (2018) Analysis of characteristics of driving behavior of commercial vehicles under big data environment. *Traffic Inf Saf* (4):42–50
20. Wu J, Fan W (2015) Research on risk assessment system of road transportation of dangerous goods. *Highw Transp Sci Technol* 32(12):6–11
21. Xiao W (2009) Research on transportation safety evaluation of automobile units. Nanjing University of Science and Technology
22. Zhu X, Lu Y, Zhou D (2006) Application of fuzzy evaluation mathematical model in road traffic safety evaluation. *Saf Environ Eng* 13(03):102–105
23. Xi B (2014) Research on comprehensive evaluation model of automobile safety. Southwest Jiaotong University
24. Xiao P (2010) Research on the safety evaluation index system of road passenger transport enterprises in Shanxi Province. Chang'an University

Vehicle Detection at Night Based on the Feature of Taillight and License Plate



Guosheng Ma, Mante Cai, Guanliang Chen, and Zhixiao Li

Abstract This paper proposes a algorithm of vehicle feature extraction and detection based on video data for night time. The color characteristics of taillights can be roughly divided into two parts no matter how far or near they are: inner ring-highlights area partial to pink and outer ring-high saturation Red areas. Through a large number of sampling statistics, this method obtains the accurate threshold range of each layer based on HSV color space. Thus, the suspected area of the inner and outer ring of tail lights can be segmented accurately and filtered preliminarily according to the shape characteristics of the tail lamp. In order to improve the detection rate and image recognition quality, the paper carry out AOI region segmentation and median filtering. After getting the suspected area of license plate, the tail light and license plate are combined to determine the rear of the vehicle. Secondly, all the connected regions of the tail lamp suspected area are paired and the confidence level of the pair is established. The confidence level is evaluated according to the characteristics of the tail lamp pair such as the horizontal height, the distance width and the symmetry centered on the license plate. According to the confidence level, whether it is qualified to pair with the license plate is determined Finally, according to the characteristics of taillight pairs, the mismatched relationship pairs are eliminated and the vehicles are identified. The experimental results show that the method can accurately detect the vehicle tail light features to identify the vehicle, and the false detection rate is low.

Keywords Vehicle detection · Taillight · License plate

G. Ma (✉) · M. Cai · G. Chen · Z. Li
School of Automobile Engineering, Harbin Institute of Technology, Weihai 264209, China
e-mail: gshma@sohu.com

Z. Li
e-mail: qq470645072@163.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,
Lecture Notes in Electrical Engineering 775,
https://doi.org/10.1007/978-981-16-5429-9_55

723

1 Introduction

Vehicle detection at night has always been a problem, mainly due to the following reasons [1–4].

- Not enough light at night. At night, the ambient light is not as good as that in daytime. It is not easy to get the details of the front vehicle contour, and the image details are lack. Generally, only the taillights are clearly visible.
- It is difficult to judge the tail light. There are many kinds of lamp forms, so we need to find a quantitative feature to describe the lamp. At the same time, there is little difference between the red lights in the light and environment.

This paper summarizes three main bases for taillights at night [5, 6].

- The brightness of the tail light is higher than the background.
- No matter the vehicle is far away or near, the basic color characteristics of the tail lamp are that the inner circle of the tail lamp area is white and pink, and the outer circle is red.
- The shape and size of tail lamp conform to certain rules.

According to these features, this paper uses the color and shape of tail lamp to extract and constrain the suspected area, and combine the license plate as an element to achieve the matching of tail lamp pairs. The research scene of the algorithm is limited to the night city road with less dense vehicles, which has certain adaptability to the environment.

2 Preprocessing Based on Taillight Color

2.1 Segmentation of Interest Area

In order to remove the interference information which is not related to detection, AOI segmentation is carried out in this paper.

The upper part of the original image contains billboards, street lights and other high-altitude environment lights, so the upper part of the area is removed; the bottom part contains irrelevant information such as front car cover reflection, so the lower part of the original image is also removed.

The specific division scope shall be determined according to the actual situation. Fix the camera behind the windshield in the car, about 1.45 m above the ground, and the shooting angle straight ahead. The video obtained is based on the flat and bumpless road surface, so the tail light and license plate display in the image focus on a narrow area in the middle. The experimental results show that the AOI region between $45.8\% \times \text{height}$ and $72.0\% \times \text{height}$ based on the image pixel height is the best. Whether the vehicle appears from the left or from the right or the turning

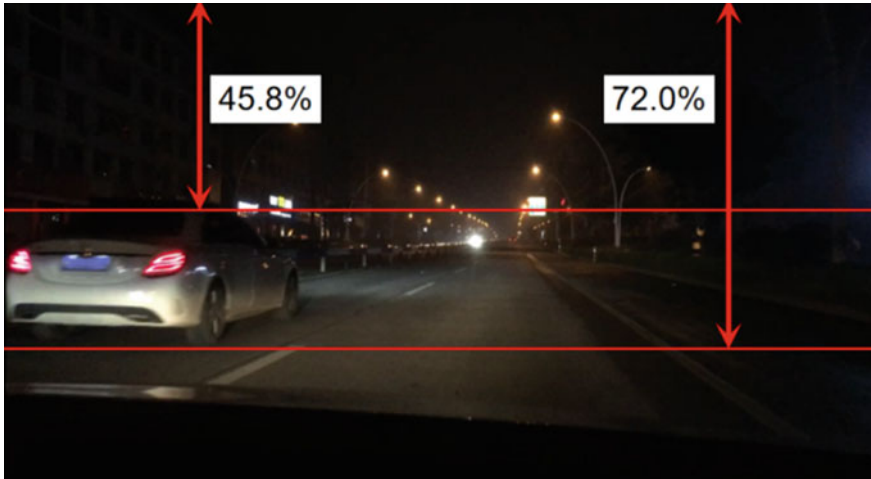


Fig. 1 Segmentation of upper and lower edge

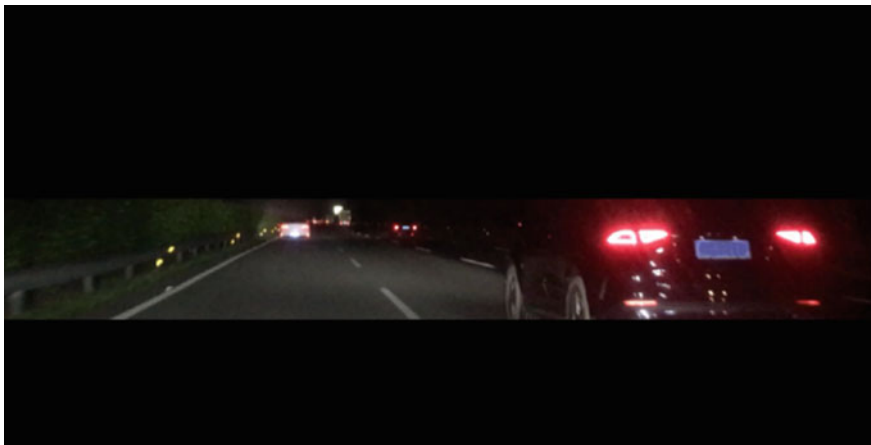


Fig. 2 Segmentation effect of AOI

vehicle, its tail lamp and license plate can be included. The scope and results of region of interest segmentation are shown in Figs. 1 and 2.

2.2 Filter Processing

According to the noise characteristics of night images, median filtering technology is suitable. This is a non-linear processing method to remove noise. Under some

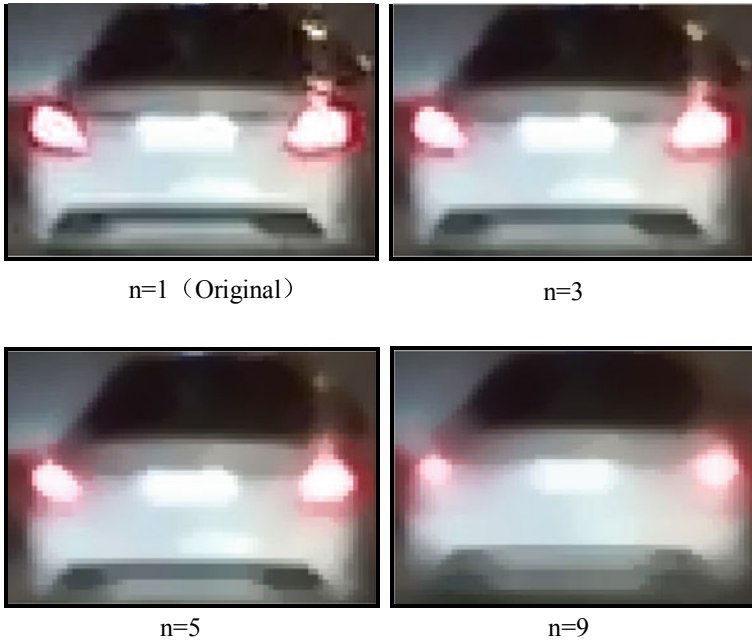


Fig. 3 Filtering effect of different n

conditions, it can not only remove noise but also protect image edge. Therefore, it can obtain better image restoration effect through median filtering [7].

In the application of image processing, the RGB image is a three-dimensional image, so the median filter needs to process three dimensions of data. The points on the edge of the image are generally irrelevant, so you don't need to consider the edge points. You can start from the second column and the second row. The processing results of the filtering method in different template sizes are shown in Fig. 3.

It can be seen that the same filtering method is not that the larger the template, the better the filtering effect. By comparison, the effect of median filter with 3×3 template is the best. With the increase of the size of template, the effect of noise elimination is enhanced, but the level of detail sharpening is reduced correspondingly, and the image becomes blurred. Thus, the loss of image detail information will destroy the accuracy of the image, and the edge will become blurred. Therefore, choosing the appropriate size of the median filter window can remove the image noise while maintaining the image accuracy to the maximum extent.

2.3 Selection of Color Space

The original image is displayed in RGB color space. The R, G and B components are represented as the proportion of red, green and blue between [0, 255], which have high affinity for computers, but do not conform to human visual characteristics. So they are often converted into other color spaces. HSV color space can be described quantitatively through three dimensions of Hue, Saturation and Value, and the selection of the judgment threshold of color will be more humanized and accurate. Therefore, this paper uses HSV color space to judge the taillight color [8–10].

3 Determination of Color Threshold

This article collected 5 videos taken by the camera in different scenes. Frames of five video are extracted and saved respectively, and then the image is preprocessed for each frame, and the car taillight and license plate pixel information in the processed image is sampled. Each video is sampled every 10 frames, every video has a total of about 700 pixels, records RGB information, and creates an EXCEL table to convert the data to HSV data. A total of about 3500 sets of data are sampled in these 5 videos. The idea of sample point collection is shown in Fig. 4.

During the acquisition process, it was found that with the distance between the car in front and the camera, the performance of the tail lights on the image is also different, but there are only two color forms, one is the inner circle of the tail light with pink highlight area, and the other is the outer circle with high saturation red area. And it should be noted that the distance will affect the respective threshold value range of the two regions. When the distance between cars is closer, the red area of the outer ring is highly saturated and brighter, and the color is relatively single, so the threshold range is also single. When the distance is farther, the color gamut of the outer ring becomes complicated due to the influence of light. The threshold range will also become complex and diverse. No matter how far the inner circle is, the color characteristics are relatively similar. The pictures are shown in Figs. 5 and 6.

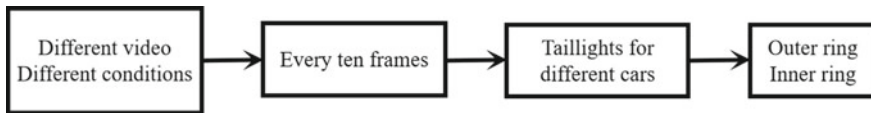
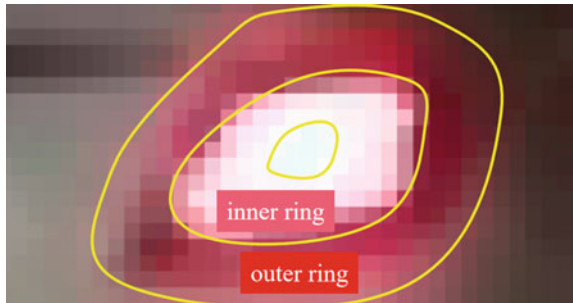


Fig. 4 The idea of sample point collection

Fig. 5 Tail lights when the vehicle is closer



Fig. 6 Tail lights when the vehicle is far away



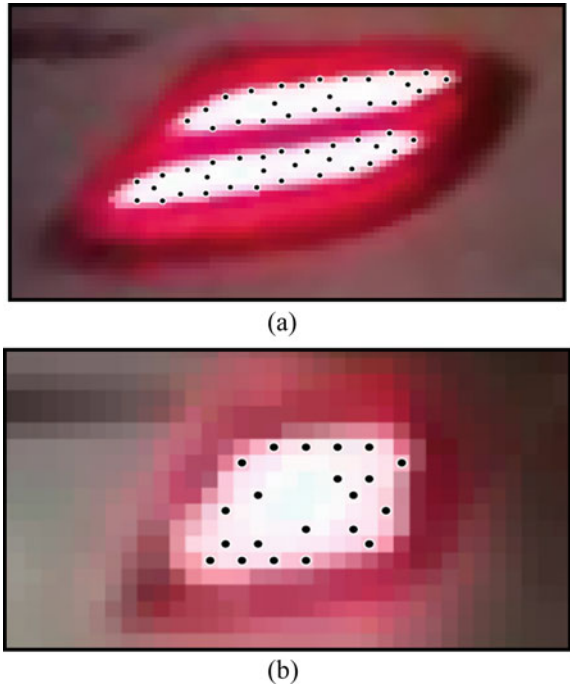
3.1 Determination of Inner Circle Threshold

The color gamut of the inner circle is a pink highlight area, the chroma is red, the brightness is higher and the saturation is lower, so it appears pink. The selection of sample points needs to pay attention to these.

- The sample points are evenly dispersed, and the points are selected in discontinuous rows or columns.
- Enough points are selected to cover and surround the inner circle area evenly.
- Avoid points with higher saturation or darker colors (to be selected when studying the outer circle, and now the selection will cause the jump of the inner circle threshold range).
- Try not to choose the highlight points with green or blue in the inner circle area, because they do not conform to the characteristics of the tail lights, causing interference to the threshold range.

Example of sample point selection in inner circle is shown in Fig. 7. After the statistical points are converted to other HSV, the points with similar data distribution range are classified into one group through processing. The unprocessed data distribution is very confusing, and it is difficult to see the law. Points are grouped into one category. It should be noted that because the display of pixels is determined by HSV at the same time, the similarity of data distribution must be based on the three-dimensional simultaneous similarity of HSV.

Fig. 7 Example of sample point selection in inner circle



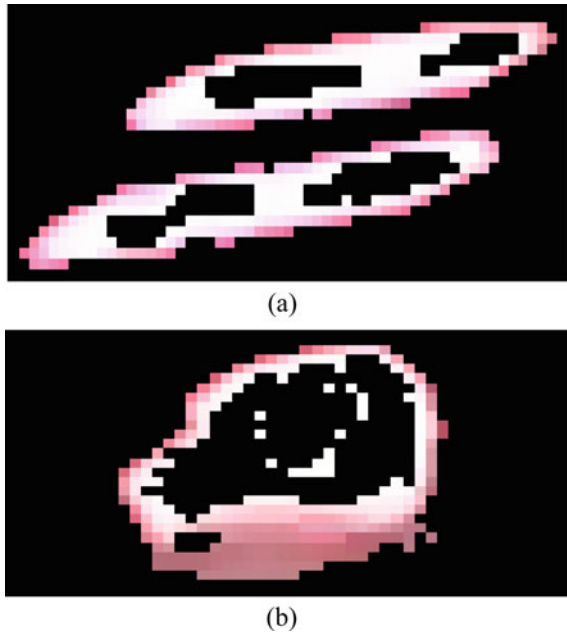
After removing some extreme values, the distribution range of each dimension of HSV is as narrow and close as possible, then the three-layer threshold range can be summarized, and shown in Table 1.

Using these two thresholds to segment the image at the same time, the experimental results obtained are shown in Fig. 8.

Table 1 The three-layer threshold range

A	Min	Max
Hue	0.9	1
Saturation	0.286	0.56
Value	0.7	1
B	Min	Max
Hue	0.88	1
Saturation	0	0.307
Value	0.835	1
C	Min	Max
Hue	0	0.267
Saturation	0	0.307
Value	0.835	1

Fig. 8 The effect of threshold segmentation

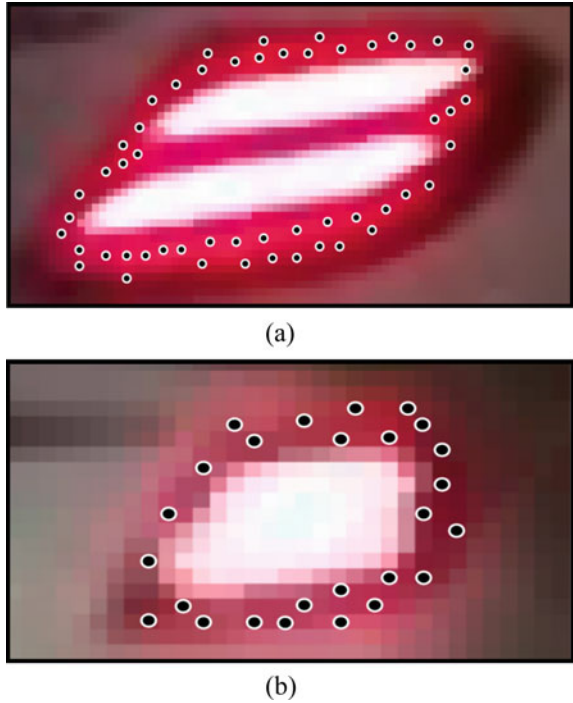


According to the experimental results of different videos, the result of A is ideal, the non-tail light area is less, and the tail light area is displayed well, but the B's and C's display result is not good. Many intolerable non-taillight parts are mostly road lights, billboard lights and other parts outside the inner and outer rings, which appear as yellow and gray highlight light sources, So the threshold of B and C needs to be corrected. Each layer of threshold has the possibility of correction.

3.2 Determination of Outer Circle Threshold

The selection of sample points of the outer ring is similar to the inner ring, and they need to be evenly distributed, and there can be enough points to surround the outer ring to shape the tail light (the outer ring of the tail light can be divided). It should be noted that the diversity of the color of the sample points will change with the distance of the outer ring, and the sample points that are purplish and brown may be part of the taillights. However, due to blurry focus, light divergence, strong light interference, etc., the colors of some outer circle pixels that are too far away from the inner circle will become similar and are not necessary to be selected. The necessary sample points, even if they are similar in color, and the sample points that determine the shape of the tail light, even if the chromaticity difference is large, should also be taken into account. Example of sample point selection in outer circle is shown in Fig. 9.

Fig. 9 Example of sample point selection in outer circle



The number of pixels in the outer circle of the tail light is more than double that of the inner circle. Similarly, the data convert into HSV.

After processing, the data can be roughly divided into two parts. If the experimental result is more ideal, it is considered to be a part.

The experimental results generated by the two thresholds at the same time are shown in Fig. 10. Similarly, there are segmentation results of non-ideal regions, so they also need to be corrected.

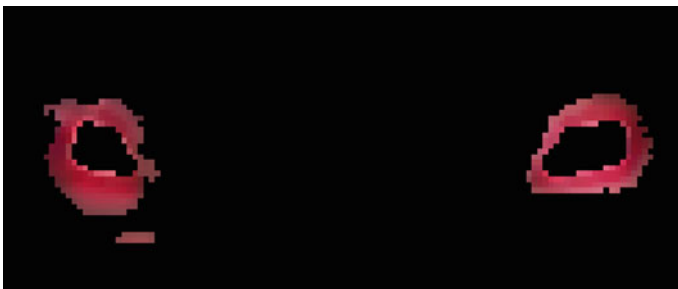


Fig. 10 The effect of threshold segmentation

3.3 Threshold Correction

Each layer of threshold application may need to be corrected in actual situations. The usual method of correction is to select the pixels of the non-tolerable non-taillight region (which causes adhesion and deforms the connected domain and affects subsequent taillight pairing) HSV data, narrow the range of the corrected layer to avoid the HSV value of the non-ideal point. For example, there are many non-ideal areas in B and C. Use the same method to obtain RGB information of non-ideal pixels and convert them. The threshold distribution of non-ideal areas is calculated and shown in Table 2.

It should be noted that when avoiding this range, the pixel value is determined jointly by the three of HSV, and they are indispensable. Therefore, it is not necessary to reduce the range of the HSV three to modify at the same time, just choose one or two to avoid, so There are various correction schemes, and the most ideal correction results among the various schemes are show in Table 3.

This constraint is only corrected by H. As a result, the tail light shows that there are more pixels in the inner circle and the circle layer is thicker, but there are also some other light sources. These light sources are similar to the pixels in the inner circle, and it is difficult to circumvent through the color threshold.

The outer circle can be corrected in the same way. So far, there are a total of five layers of thresholds. Each layer must be separate and cannot be combined. For example, point $H = 0.95, S = 0.47, V = 0.95$, it may not belong to any of the five layers, but it may exist after the merger range, so merging will expand the range of color segmentation and misjudge many non-ideal areas. Finally, the five-layer modified threshold is used to perform color segmentation to obtain the tail light area I_taillight (I_tl for short) as shown in Fig. 11. And the effect is nice.

Table 2 The threshold distribution of non-ideal areas

Non-ideal	Min	Max
Hue	0	0.167
Saturation	0	0.13
Value	0.875	1

Table 3 The most ideal correction results

B (Corrected)	Min	Max
Hue	0.88	1
Saturation	0	0.307
Value	0.835	1
C (Corrected)	Min	Max
Hue	0.17	0.267
Saturation	0	0.307
Value	0.835	1



Fig. 11 Experimental effect of threshold segmentation

4 Generation of Suspected Tail Light

The algorithm filters out the suspected tail light area by threshold value, and then binaries the filtered points [11]. The binarization method is ‘Either black or white’, that is, as long as the point in the figure is not black background ($RGB = 0$), it takes as 1. After that the connected region is calculated and labeled [2, 11], and the external rectangle is extracted for frame selection. Finally, the suspected area, including the taillight, is extracted from the screen, and the suspected area is preliminarily obtained, as shown in Fig. 12.

As can be seen from the figure, there are a large number of suspected tail lamp areas. In order to remove some objects that are obviously not lamps, the algorithm filters the shape of the external rectangles of these areas roughly. The types of areas to be removed are: areas with too small area; areas with too small height and width; areas with too large or too small aspect ratio, because the tail lamp will not be thin or flat. After the initial filtering based on the shape of the external rectangle, most

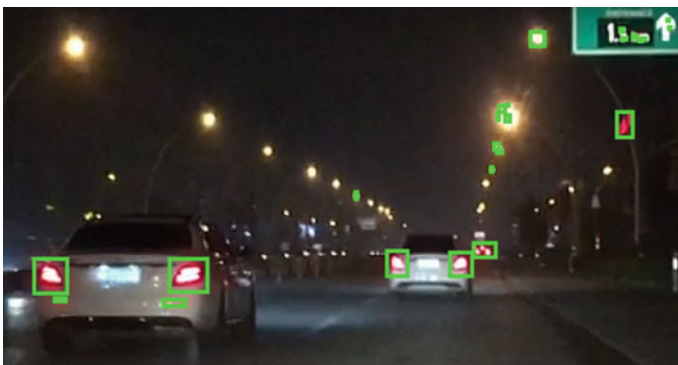


Fig. 12 Results of threshold segmentation

of the interference light sources are excluded, and the suspected lamp area is further extracted. Binary figure after shape filtering is shown in Fig. 13.

Then label the filtered tail lamp binary graph. All points of the nth connected domain are assigned to n from 1, which realizes the labeling of connected domain. The marking process of the fifth and sixth connected domains in the binary graph is shown in Fig. 14.

The pixel value of the white pixel was originally 1, and was given values 5 and 6 respectively after the labeling process. Figure 15 is part of the matrix intercepted by the fifth and sixth connected domains. It can be seen that the values of all positions of the connected domains have been replaced by their respective serial numbers from 1.

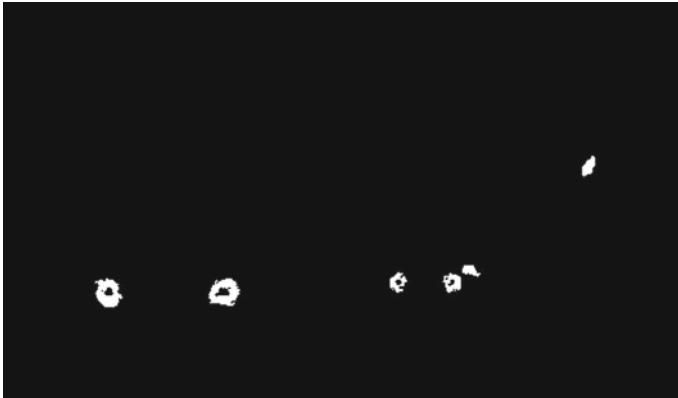


Fig. 13 Binary figure after shape filtering

Fig. 14 The fifth and sixth connected domain in binary graphs



	695	696	697	698	699	700	701	702	703	704	705	706	707
362	0	0	0	0	0	0	0	0	0	0	0	0	0
363	0	0	0	0	0	0	0	0	0	0	0	0	0
364	0	0	0	0	0	0	0	0	0	0	0	0	0
365	0	0	0	0	0	0	0	0	0	0	0	0	0
366	0	0	0	0	0	0	0	0	0	0	0	0	0
367	0	0	0	0	0	0	0	0	0	0	6	6	6
368	0	0	0	0	0	0	0	0	0	0	6	6	6
369	0	0	0	0	0	0	0	0	6	6	6	6	6
370	0	0	0	0	0	0	0	0	6	6	6	6	6
371	0	0	0	0	0	0	0	0	6	6	6	6	6
372	0	0	0	0	0	0	0	0	6	6	6	6	6
373	5	0	0	0	0	0	0	0	0	0	6	6	0
374	5	5	5	5	5	0	0	0	0	0	0	0	0
375	5	5	5	5	5	5	0	0	0	0	0	0	0
376	5	5	5	5	5	5	5	0	0	0	0	0	0
377	5	5	5	5	5	5	5	0	0	0	0	0	0
378	0	0	0	5	5	5	5	0	0	0	0	0	0
379	0	0	0	5	5	5	5	0	0	0	0	0	0
380	0	0	0	5	5	5	5	0	0	0	0	0	0
381	0	0	5	5	5	5	5	0	0	0	0	0	0
382	5	5	5	5	5	5	5	0	0	0	0	0	0
383	5	5	5	5	5	5	5	0	0	0	0	0	0
384	5	5	5	5	5	5	5	0	0	0	0	0	0
385	5	5	0	5	0	5	0	0	0	0	0	0	0
386	5	5	0	5	0	0	0	0	0	0	0	0	0

Fig. 15 Connected domain matrix after labeling

This realizes the label of the connected domain, which is convenient for the later matching process. After labeling the connected region, we can easily obtain the information such as rectangle, area, center and so on. The definition of information is shown in Fig. 16. The required information of this algorithm is: the abscissa Tx_axis and ordinate Ty_axis of the upper left corner of the connected region, width of the rectangle Twidth and height Theight. From this, the coordinates Tx and Ty of the center point of the rectangle can be calculated.

The center coordinates of the circumscribed rectangle of the ith connected domain are

$$Tx(i) = Tx_axis(i) + Twidth(i)/2 \tag{1}$$

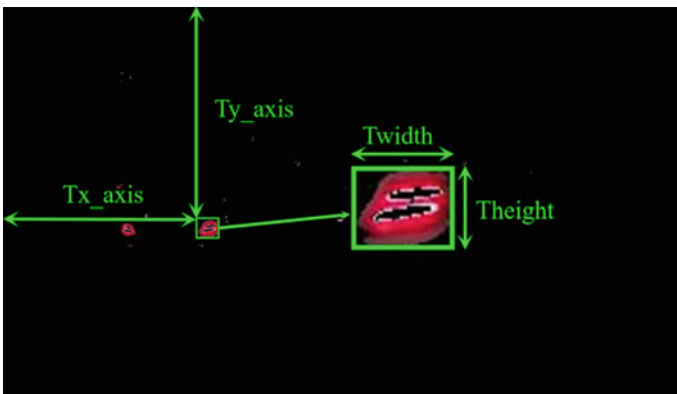


Fig. 16 The definition of information

$$Ty(i) = Ty_axis(i) + Theight(i)/2 \quad (2)$$

So far, the generation of tail lamp suspected area is over.

5 Generation of Auspected License Plate

Considering the factors of license plate, the matching accuracy and efficiency can be improved. The algorithm recognizes the left and right tail lights, and then recognizes the license plate. Using the association information of these three areas, a car can be determined.

As for the license plate and tail lamp, the display of different distance on the image will be different. When the distance is relatively close, the license plate display is clear, and the blue background and its boundary, even the font outline, are relatively clear. At a little distance (10 m), the license plate begins to be white and has no font features. The white area accounts for a large part, but there is still an obvious blue part, and the license plate contour boundary is still easy to identify. A little further away (20 m), the license plate is completely white, the outline is still clear, there is no visible blue feature, the overall white regular rectangle.

In the process of extracting the blue license plate, the method is the same as that of tail lamp extraction. A large number of RGB pixel values are sampled and converted into HSV form, about 700 pixel points, and five layer thresholds are summarized. For example, layer D and layer E are shown in Table 4, and the connected domain I_license plate (I_lp for short) of one of the segmentation results is shown in Fig. 17.

The corresponding display effect of D and E layers is as shown in Fig. 18.

The practical application results in the license plate area I_lp, the result is more ideal, but there is not much space for correction. The color of the non ideal area is indeed the necessary color of the license plate area. Then, it is also necessary to filter the license plate area I_lp based on shape. The main filtering ideas are as follows.

Table 4 Five layer thresholds are summarized

D	Min	Max
Hue	0.49	0.556
Saturation	0.1	0.4
Value	0.8	1
E	Min	Max
Hue	0.5	0.65
Saturation	0.1	0.35
Value	0.5	0.81



Fig. 17 Relationship between display effect



Fig. 18 Display effect of license plate

- The results of threshold segmentation in different distances should be classified separately. The areas within 10 m are classified as i_lp_0 , those within 10–20 m are classified as i_lp_10 , and those beyond 20 m are classified as i_lp_20 . For example, the areas divided by D and E are i_lp_10 . Then, the following operations are carried out respectively.
- Removing the connected region with aspect ratio not conforming to the characteristics of license plate.
- The length width ratio of the external rectangle of the license plate area recognized by an efficient algorithm will not be too large or too small, which is close to the length width ratio of the actual license plate. Definition of width and height of license plate is shown in Fig. 19. According to the experimental statistics, in the algorithm of this chapter, the aspect ratio of I_lp falls within $[1, 3.5]$.

For the i th connected domain, M = width of the external rectangle, N = height of the external rectangle, length width ratio $R = M/N$.

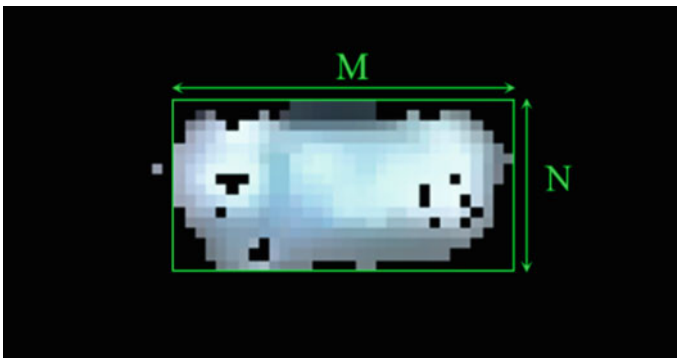


Fig. 19 Definition of width and height of license plate

$$\text{if } R > 3.5 \text{ || } R < 1, \text{ then } I_{lp}(i) = 0 \quad (3)$$

That is to erase the i th connected domain.

- Remove the connected region whose area is too small for the external rectangle.

A highly efficient algorithm, which recognizes the license plate connection domain, pixel points should be dense enough, rather than full of holes. The proportion of the area of the connected domain to the outside rectangle is not too low. In this chapter, the area ratio is within $[0.53, 1]$, and 0.53 is the minimum value that can be tolerated.

For the i th connected domain, M = width of the external rectangle, N = height of the external rectangle, and the area of the connected domain is S , $RR = S/(M \times N)$.

$$\text{if } RR < 0.53, \text{ then } I_{lp}(i) = 0 \quad (4)$$

That is to erase the i th connected domain.

The label process of the filtered I_{lp} is the same as that of the tail lamp, and the center coordinates of the rectangle outside the connected region of I_{lp} are calculated.

6 Generation of Tail Lamp Confidence

In the tail light I_{tl} area divided by color, each connected region can match at most with another connected region. How can we judge which connected region i matches with the other connected regions with the highest probability of success? In this algorithm, the idea of I_{tl} confidence is introduced.

6.1 Obtaining Confidence

The number of connected domains in the filtered I_{tl} is $numT$, and a zero matrix of $numT \times numT$ is created.

Taking the data above the main diagonal, the number in row I , column J ($J > I$) indicates the confidence level of the relationship between the i th connected domain and the j th connected domain. Then set three conditions:

- Symmetry

The abscissa of the center of the circumscribed rectangle of the i th I_{lp} is $Lx(i)$, the abscissa of the center of the circumscribed rectangle of the i th I_{tl} is $Tx(i)$, and the set of the abscissa of all the circumscribed rectangles of the i th I_{tl} is Tx . If the two taillights are matched successfully, there must be at least one license plate I_{lp} in the middle. The distance between them and $Lx(i)$ of I_{lp} is less than a certain empirical value Δ (the empirical value is calibrated by the experiment) and the direction is opposite.

$$|(Tx(t) + Tx)/2 - Lx(i)| < \Delta \quad (5)$$

$$(Tx(t) - Lx(i)) * (Tx - Lx(i)) < 0 \quad (6)$$

If the conditions of this layer are met in TX, the confidence level will be increased by one level. It should be noted that there may be more than one I_lp connected domain in the middle of a pair of taillights. For example, the I_lp satisfies the condition, and the $I + 1$ th I_lp may also satisfy the condition. Therefore, if the frequency is greater than 1, frequency is 1, algorithm is needed to limit it.

- Height difference

The horizontal height of the two taillights of the same car must not differ too much.

$$|Ty(i) - Ty| < \Delta_{\text{height}} \quad (7)$$

If the condition of this layer is satisfied, the confidence level will be increased one level.

- Horizontal distance

The horizontal distance between the two taillights of the same vehicle must not be too much different. Based on the transverse coordinate $Ty(I)$ of the outer rectangular center of the i th I_tl , find all the satisfied conditions in Ty :

$$|Tx(i) - Tx| < \Delta_{\text{width}} \quad (8)$$

If the condition of this layer is satisfied, the confidence level will be increased one level.

Finally, the tail lamp relationship pair with full confidence level is qualified to be combined with the license plate and successfully matched. The algorithm will find the connected region I_lp in the middle of the tail lamp pair, and frame three as the final recognition results.

6.2 Removal of Misidentification

However, there are still many mismatches in this confidence system, which can be divided into the following categories.

- One for many

One tail lamp can only be matched with another under normal conditions, and it is impossible to match multiple tail lamps. In this algorithm, the basic reason for this situation is that there are multiple relationship pairs with full scores in the same row



Fig. 20 One to many mismatches

of the frequency matrix. So we need to judge more conditions to further limit the confidence.

As shown in Fig. 20 on the right, if there is one car with many cars, assume that the j th and $j + 1$ th columns in the i th row are full marks at the same time, then it is necessary to determine which relationship pair is invalid. According to the first condition of confidence, if the full score can be obtained, then there must be I_{lp} in the middle of the taillight pair. Because the interference of the two correct tail lamp pairs is extremely low, the two tail lamps that are matched together should be the two closest to the I_{lp} .

According to this principle, other distant relation pairs can be eliminated, and $TX(I)$, $TX(J)$, $TX(j + 1)$ can be taken as the difference and order with the horizontal coordinate LX of I OULP center respectively:

$$D = [|Tx(i) - Lx|, |Tx(j) - Lx|, |Tx(j + 1) - Lx|] \tag{9}$$

$$[DR, index] = \text{sort}(D) \tag{10}$$

DR is the matrix after sorting from small to large, $index$ is the sequence number before corresponding sorting, the first two corresponding frequencies are reserved, and other frequencies are eliminated, so that one to many mismatches can be eliminated.

- Suspected light source interference

In the external environment, there are often interference light sources that are similar to the tail lamp in color and shape, such as billboards. They are separated from the tail lamp by the partial red field together with the tail lamp, and there may also be I_{lp} satisfying the conditions in the middle, so there is a possibility of mismatch. Mismatch of interference light source^{1.25} is shown in Fig. 21. Different from the tail lamp pair, the area of the two connected areas of the interference light source is often smaller and far away, so the principle that the larger the distance between the

Fig. 21 Mismatch of interference light source1.25



two tail lamps is, the larger the tail lamp area should be can be used to eliminate the mismatched lamp pair.

Definition of center distance and connected area is shown in Fig. 22. The connected area of the two taillights is $tstats(1).Area$ and $tstats(2).Area$, and the center distance is set as X . Take:

$$Actualarea = \max(Tstats(1), Tstats(2)) \tag{11}$$

$$Y = \min(Tstats(1), Tstats(2)) \times (1 - 0.1) \tag{12}$$

In the tail lamp image segmented by this algorithm, for different x statistic Y values as far as possible, it should be noted that the smaller the Actual-area is, the less obvious the distance between vehicles in different lanes with the camera is (When x is getting bigger). The smaller the Actual-area is, the less obvious it is. On the contrary, the larger the X is, the more obvious the Actual-area is, so the trend

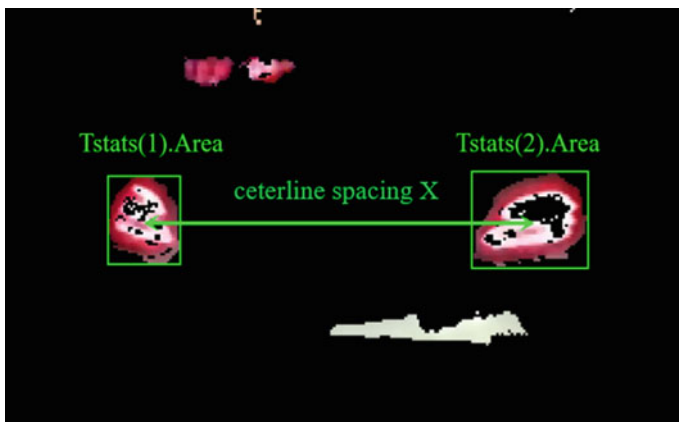


Fig. 22 Definition of center distance and connected area

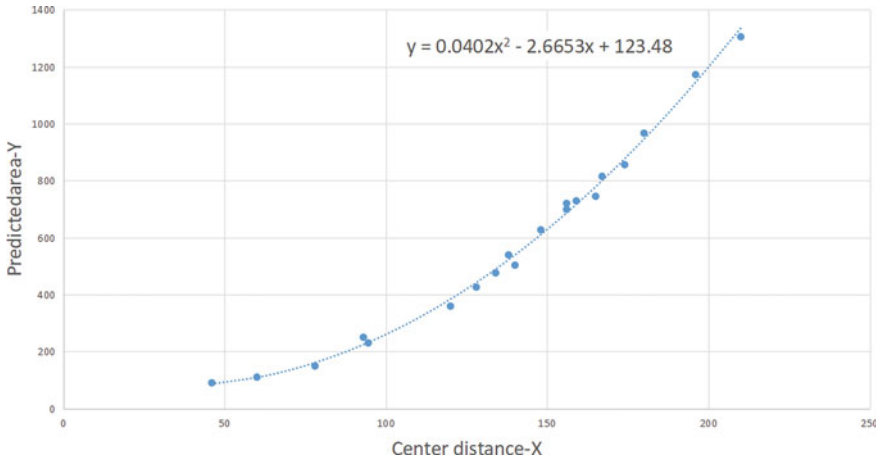


Fig. 23 Relationship between predicted-area and X

line should be a nonlinear polynomial. In this paper, about 20 groups of X–Y data under different conditions are counted, and the number of statistics within [100200] is more, because the probability of error in this part is higher.

As shown in Fig. 23, Y represents the predicted-area, and the judgment relationship is:

$$\text{if Actual-area} < \text{Predicted-area, then frequency} = 0$$

The largest connected area of two suspected taillight pairs is smaller than predicted area, so it is a non ideal relationship pair.

- Resolve adhesion

When the red area in the image is too close, the algorithm is likely to identify them as the same tail lamp suspected area, and through layer by layer filtering, all conditions are met, and finally the matching is successful, as shown in Fig. 24. The solution is simple. For example, if the *i*th and *j*th match successfully and there is adhesion, it is only necessary to replace the length and width of the rectangle outside the larger connected domain with the length and width of another connected domain.

7 Final Experimental Results

With the check of the algorithm of eliminating false recognition, there are very times of missing or false detection.

In this paper, VC++ is used for experiment. Experimental results of the algorithm is shown in Fig. 25. The experiment results show that the recognition effect is good.

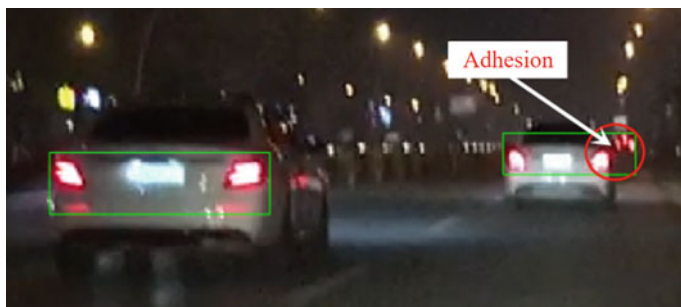


Fig. 24 Adhesion

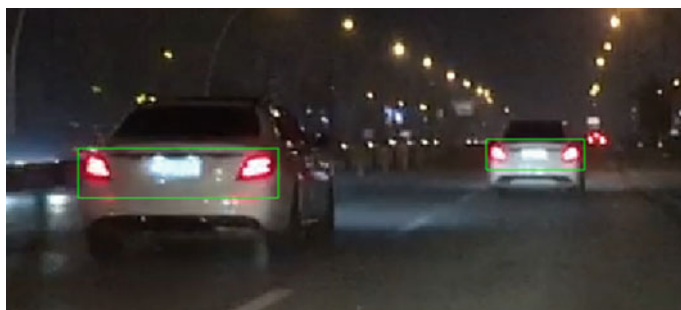


Fig. 25 Experimental results of the algorithm

The algorithm can accurately identify and frame the area including tail lamp and license plate. The detection and recognition rate of the algorithm on urban main roads and highways are shown in Table 5.

There are many non taillight interference light sources on urban main roads. There are 43 frames of 582 frames of video with false detection (including false detection and missing detection), which shows that the algorithm can effectively eliminate the influence of interference light source, and the situation of false detection is mostly that the taillight area is interfered by the front headlights, and the red area is covered by strong light, so it cannot be detected. Interference by headlights is shown in Fig. 26.

To summarize, in the case of less vehicles and less environmental interference, the recognition effect of the algorithm is very considerable.

Table 5 The detection and recognition rate of the algorithm

	Total frame number	False detection frame number	Recognition rate (%)
City	582	43	92.6



Fig. 26 Interference by headlights

References

1. Sun Z, Bebis G, Miller R (2004) On-road vehicle detection using optical sensors: a review. In: The international IEEE conference on intelligent transportation systems, 2004 proceedings. IEEE, Washington, DC, pp 585–590
2. Yu LY, Guo YL (2019) Method of preceding vehicle detection and tracking at night based on taillights. *J Tianjin Polytech Univ* 38(1):61–68
3. Liu ZY, Ye Q, Li F et al (2010) Taillight detection algorithm based on four thresholds of brightness and color. *Comput Eng* 36(21):202–203, 206 (in Chinese)
4. Tian XX, Wang JS (2016) Application of finite state machine in night time vehicle detection. *J Shijiazhuang Tiedao Univ (Nat Sci)* 29(4):95–99
5. Lin YC, Lin CC, Chen LT, Chen CK (2011) Adaptive IPM-based lane filtering for night forward vehicle detection. In: IEEE conference on industrial electronics and applications. ICIEA, Beijing
6. Xiao ZT, Wang Y, Geng L, Zhang F (2013) Preceding vehicle detecting method for night based on geometry and position features of blobs. *J Hebei Univ Technol* 42(5):13–18 (in Chinese)
7. Xu CC (2019) Denoising algorithm based on image histogram. *Comput Internet* 45(16):34–35
8. O'Malley R, Jones E, Glavin M (2010) Rear-lamp vehicle detection and tracking in low-exposure color video for night conditions. IEEE
9. Bao QL (2010) Colorful image segmentation based on HSV. *Softw Guide* 09(07):171–172
10. Qi QH, Chen QX (2012) Vehicle detection based on two-way multilane at night. *Commun Technol* 45(10):58–60
11. Tian Y (2009) Digital image processing and analysis. Huazhong University of Science and Technology Press, Wuhan

Spatial Integration Based on Big Data—A Case Study on the Integration of Nanjing and Ma’anshan



Xiaofeng Lou and Hong Zhang

Abstract By carrying out a case study on the integrated urban development of the cities of Nanjing and Ma’anshan in China, we developed a method to analyze the integration characteristics and assess the integration of adjacent areas based on multi-dimensional analysis of big data on population and industry. We investigated the spatial distributions, connections, and the integration of urban elements such as population, industry, and transportation. We evaluated the degree of integration of the two cities based on multiple attributes. The findings will guide the integration planning of key areas for the next phase of urban development and provide new ideas for research on integrated development.

Keywords Integration · Big data · Ma’anshan · Nanjing

1 Introduction

In February 2019, the National Development and Reform Commission of the People’s Republic of China issued “Guidance on Cultivating and Developing Modern Metropolitan Areas.” In December 2019, the Central Committee of the Communist Party of China and the State Council of the People’s Republic of China issued “Outline of the Regional Integration Development Plan of the Yangtze River Delta” to promote regional integration and high-quality development. Integrated planning is pursued when urbanization reaches advanced levels. Because the Yangtze River Delta region is considered the economic and technological center of China, much attention has been paid to the city clusters integrated in the region [1].

Under the principles of ecological civilization, the standards of the quality and utilization efficiency of the national space are higher. Hence, developed areas change gradually to become ecologically friendly, and medium- and small-sized cities adjacent to these developed areas benefit from these changes as they seek industrial

X. Lou (✉) · H. Zhang

School of Architecture, Southeast University, Sipailou 2, Nanjing 210096, China

upgrading [2]. Therefore, production and daily activities are divided and coordinated among different cities and the collaboration between governments and non-governmental organizations becomes more frequent [3]. It is expected that the integration of adjacent areas is more likely to take place in the future.

As information and transportation networks develop and improve, various elements flow more smoothly and frequently between cities. Inseparable connections are formed for some cities, which are close to each other, and spatial integration becomes inevitable [4]. While spatial fusion can be observed and perceived, the flow of urban elements such as population, capital, and materials cannot be easily recognized or quantified. By adopting big data analysis, the flow and connections of these invisible elements between different spatial units can be discovered, and the characteristics of these elements can be quantitatively examined. The findings of big data analysis can facilitate integrated planning.

1.1 Current Development of Integration

Integrated planning is coordinating the urban and rural construction in a region across the spatial boundaries of different elements from a wider perspective and establishing unified development strategies. A key step in integration research is to determine the characteristics of the connections between cities. Unlike traditional spatial planning that focuses on the internal organization of a city, integrated planning is considered “unconventional” planning, which emphasizes the connection networks and systems within city clusters and highlights the roles of the leading cities and rational division of labor among cities. Integrated planning focuses on the reasonable distribution and aggregation efficiency of elements such as population and employment. In recent years, some researchers have examined the evolution of spatial integration in metropolitan areas in China and Europe. In the twenty-first century, China is expected to be the first country to have large metropolitan areas after Europe [5].

Many researchers have studied the integrated development in different regions in China [6, 7] by analyzing the implications, characteristics, mechanisms, and modes of integration [8–10]. To achieve higher accuracy in such studies, researchers employed quantitative methods to analyze the network relationships within city clusters. Li and Zhang [11] investigated the spatial structure of the Yangtze River Delta according to the distributions of city sizes, spatial forms, and connection networks. Tang et al. [12] and Wu et al. [13] measured the strengths of the economic connection between cities to examine the level and spatial pattern of the integration in the region. Chen and Wang [14] used different dimensions (density, distance, functional division, and connection) to establish indicators to evaluate regional integration and quantitatively analyzed the spatial pattern of integration in the Yangtze River Delta.

Most measurement methods are based on local annual survey data, which are mainly static cross-sectional data such as regional gross domestic product (GDP) and population size. These data are spatially too coarse to describe the structural

characteristics of smaller units. In terms of regional connection, regional macroscopic data are independent and mostly descriptive. Thus, the evaluation of a direct connection is not possible using this type of data. In terms of timeliness, the acquisition frequencies of these data are low. Therefore, the results of those methods show some lags and are not applicable to rapid economic development and elemental flow in the region of interest.

1.2 New Trends in Integration Research

Recent developments in big data analysis techniques have introduced big data to the field of planning. Niu and Li [15] utilized mobile phone signal data in 16 cities in the Yangtze River Delta to study the inter-city commuting activities in Shanghai and its surrounding cities. Similarly, Li and Chen [16] adopted mobile communication data to quantitatively analyze the commuting activities between workplaces and residences as well as the attractiveness of public service centers. Li and Chen [16] also determined the fundamentals of urban cohesion required for integration. Based on the special planning content involved in city master plans, Pang et al. [17] analyzed big data from various sources to form an integrated system to assess the implementation of planning using big data. Zhao et al. [18] examined the development trends and evolution characteristics of spatial integration in the Yangtze River Delta using remote sensing data.

To date, there has been limited research exploring the application of big data to integration research. In the existing research, only a few dimensions have been examined, and multi-source data have not been used in integration planning.

2 Technical Framework for Analysis and Indicator Construction

2.1 Technical Framework for Analysis

To address the shortcomings of the existing studies, we fully utilized big data from different sources in this paper. Traditional survey data, open data available on the Internet, and commercial data were analyzed. Then the spatial distributions of different elements were obtained in Nanjing and Ma'anshan.

In this study, the data were analyzed at the regional and city levels to reveal the characteristics of integration, develop regional integration strategies, and provide guidance for the development of key areas (Fig. 1).

At the city level, the population and enterprise data trajectories were studied to analyze the population and industry distribution in Nanjing and Ma'anshan, as well as

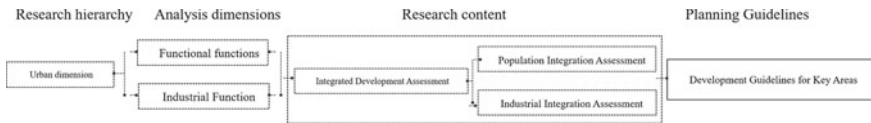


Fig. 1 Assessment of level of city. *Data Source* Qixinbao Enterprise Data

their city functions as workplaces and residences. The current state of integration was evaluated to provide guidance for planning the key areas in Nanjing and Ma’anshan.

2.2 Analysis Dimensions and Quantitative Indicators

According to the content listed in the research framework (Fig. 1), data availability, and research feasibility, several indicators covering five dimensions were used for the quantitative evaluation of integration levels (Table 1).

3 Comparison of Integration Relationships at the City Level

3.1 Activities Trajectory Integration Assessment Based on Mobile Phone Signaling Data

3.1.1 Day and Night Population Analysis

Based on the mobile phone signaling data, the cohort model was adopted to correct and allocate the total population in the region. A population distribution mesh consisting of 250 m × 250 m grids was constructed to include various data including day and night population and activity trajectories. The characteristics of population integration in the region were examined (Fig. 2).

The day and night population and the day/night differences in the population of Nanjing and Ma’anshan were subjected to kernel density analysis. The results were compared to the center-oriented planning system. Nanjing and Ma’anshan have formed a distinct structure consisting of primary and secondary core cities. Among near-region integration areas, Wujiang and Bowang have been developing more slowly. No significant aggregation of the population was observed.

Table 1 List of indicators to quantify integration

Level	Quantitative dimension	Quantitative indicator	Data source	Methodology and content
City level	Workplace and residence functions	Day/night differences in population	Enterprise data until October 2019 were purchased from Qixinbao. A total of 306,060 data points on all enterprises in Ma'anshan and 2,155,587 data points on those in Nanjing were purchased. In particular, 181,356 and 1,431,631 of these enterprises are still operating in Ma'anshan and Nanjing, respectively	Day/night difference in population = the population during the day—that at night If the value is positive, the population is greater during the day. The people mainly stay in the city for work. Contrarily, if the population at night is greater, the people mainly live in the city
		Inter-city employment connections		Inter-city employment connection = inter-city employment/total employment of the city (normalized)
	Industrial functions	Levels of industry concentration		Level of industry concentration = the number of enterprises in a certain industry/the total number of enterprises in all industries It reflects the level of concentration of an industry in an area
		Levels of industry aggregation		Level of industry aggregation = the number of enterprises in an area/that in a region It helps to determine the level of industry aggregation in an area

(continued)

Table 1 (continued)

Level	Quantitative dimension	Quantitative indicator	Data source	Methodology and content
		Ratios of registered enterprises to operating enterprises		Ratio of registered enterprises to operating enterprises = the number of registered enterprises/that of operating enterprises It reflects the characteristics of registration and operation environments. If ratio > 1, registration environment is better than the operation one. Registered economy is advantageous, but the operation environment has to be improved. If ratio < 1, the operation environment is better. The business environment has to be improved to attract enterprises to register in the area of interest

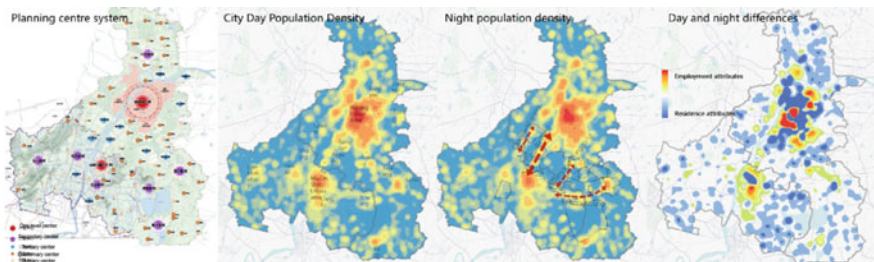


Fig. 2 Day and night population analysis of Nanjing and Ma'anshan

Table 2 Activity trajectory analysis between Nanjing and Ma'anshan

Nanjing - Ma On Shan Cross-City Link	Percentage of contact
Jiangning Street - Ma On Shan City Centre	0.00140205
Downtown Nanjing (5 km) - Downtown Ma On Shan	0.00063061
Cross Creek Street - Ma On Shan Town Centre	0.0005327
Shiqiu Town - Ma On Shan Town Centre	0.00047773
Shiqiu Town - Bowang Town	0.00038735
Tin Chung Kiu Street - Ma On Shan Town Centre	0.00032359
Nanjing Main City South - Ma On Shan City Centre	0.00031154
Lukou Street - Ma On Shan Town Centre	0.00027664
Nanjing Main City North - Ma On Shan City Centre	0.00024602
Nanjing Main City West - Ma On Shan City Centre	0.00023572
Hengxi Street - Danyang Town	0.00019807
Hengxi Street - Gushang Town	0.0001682
Jiangning Street - Yintang Town	0.00015715
Valley Street - Ma On Shan Town Centre	0.00015288
Hengxi Street - Yintang Town	0.00014836

3.1.2 Activity Trajectory Analysis

Inter-city activity trajectory data were also extracted. Kernel density analysis was conducted at the respective cities of the workplace and residence of the residents who have to commute. High connectivity is noted at the city centers of Ma'anshan and Nanjing and the streets and towns of the adjacent areas (Table 2; Fig. 3). For example, higher connectivity is observed in Nanjing-Ma'anshan near-region integration areas, namely Bowang-Shiqiu and Danyang-Hengxi.

3.2 Industrial Integration Assessment Based on Enterprise Data

3.2.1 Levels of Industry Concentration

The percentages of enterprises providing general services in Ma'anshan are smaller than those in Nanjing (Table 3). Yet, enterprises of traditional industries, such as

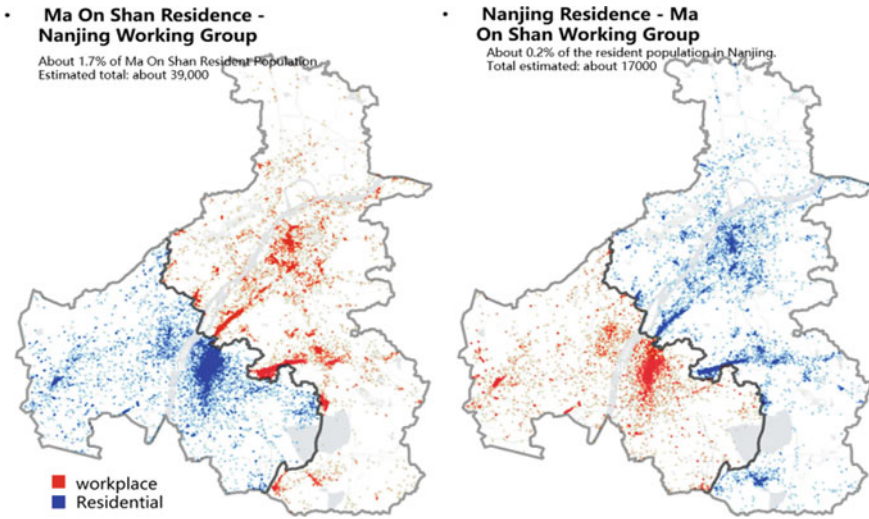


Fig. 3 Behavioral characteristics of cross-city behavior. *Data Source* Inter-city Human Trace Data for Unicom Operators

Table 3 Proportions of different types of enterprises in the two cities

Industries in Nanjing	Percentage (%)	Rank	Ma On Shan	Percentage (%)	Rank
Wholesale and retail trade	42.05	1	Wholesale and retail trade	29.62	1
Leasing and business services	18.07	2	Leasing and business services	12.07	3
Accommodation and catering	7.03	3	Accommodation and catering	1.72	11
Manufacturing	6.59	4	Manufacturing	19.09	2
Residential services, repairs and other services	5.17	5	Residential services, repairs and other services	2.24	9
Scientific research and technology services	4.34	6	Scientific research and technology services	6.63	5
Construction industry	4.23	7	Construction industry	8.65	4
Information transmission, software and information technology services	3.18	8	Information transmission, software and information technology services	2.16	10
Transport, warehousing and postal services	2.89	9	Transport, warehousing and postal services	3.51	8
Culture, sports and entertainment	2.28	10	Culture, sports and entertainment	1.39	12
Real estate	1.48	11	Real estate	3.67	7
Agriculture, forestry, animal husbandry and fisheries	1.26	12	Agriculture, forestry, animal husbandry and fisheries	6.63	6
Education	0.40	13	Education	0.11	18
Health and social work	0.30	14	Health and social work	0.23	17
Finance	0.30	15	Finance	1.16	13
Water, Environment and Public Facilities Management	0.27	16	Water, Environment and Public Facilities Management	0.39	15
Mining	0.09	17	Mining	0.30	16
Electricity, heat, gas and water production and supply	0.07	18	Electricity, heat, gas and water production and supply	0.46	14

manufacturing, construction, and real estate, as well as new industries, namely scientific research, account for higher percentages of the total number of enterprises in Ma’anshan.

3.2.2 Levels of Industry Aggregation

Kernel density analysis was conducted for all enterprises. According to the year of registration, the distribution of enterprises and its evolution over time can be analyzed. The distributions of enterprises in 2000, 2010, and 2019 reveal that Nanjing-Ma'anshan enterprises have increased their concentration along axes (Fig. 4). In Cihu-Binjiang, Qiaolin-Wujiang, Konggang-Danyang, and Lishui-Bowang, cross-region spatial distribution connections of enterprises have become closer over time.

Subsequently, kernel density analysis was performed on the industries, which all these enterprises belong to. The manufacturing and transportation industries and scientific research have shown continuous development trends over large areas (Fig. 5).

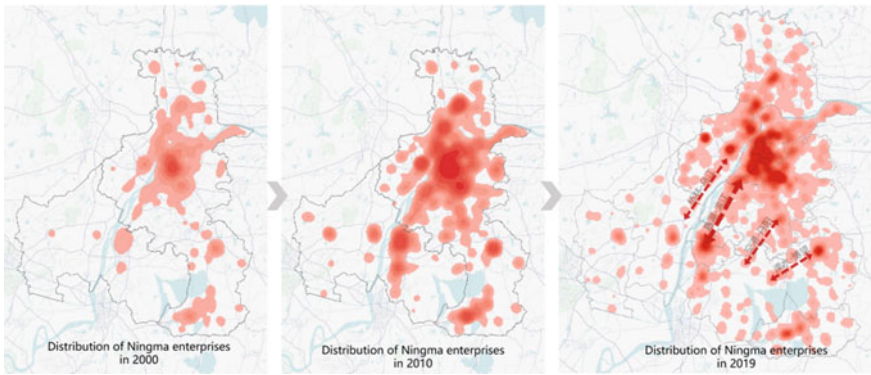


Fig. 4 Enterprise spatial evolution and industry aggregation features. *Data Source* Qixinbao Enterprise Data

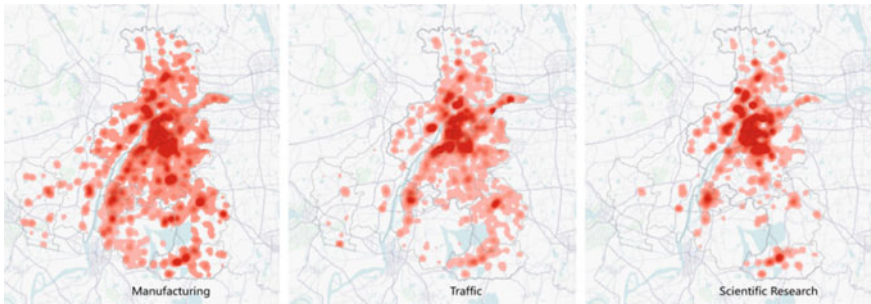


Fig. 5 Map of spatial link development industry distribution. *Data Source* Qixinbao Enterprise Data

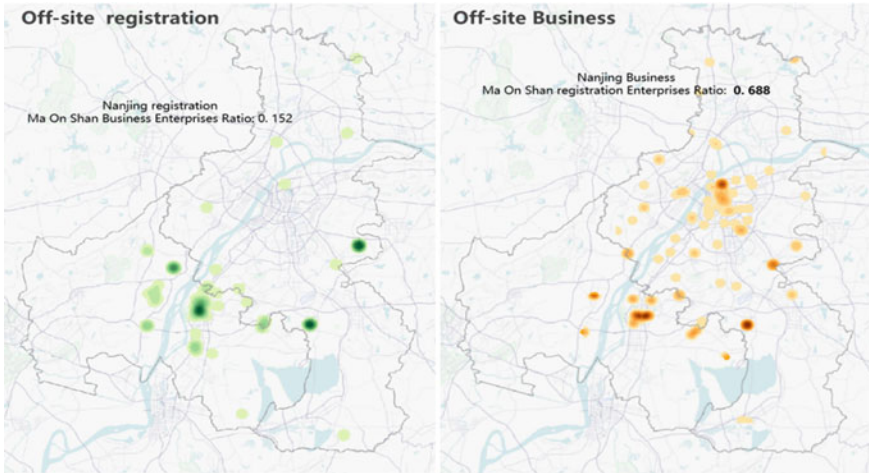


Fig. 6 Register and running of enterprises in the other city. *Data Source* Qixinbao Enterprise Data

3.2.3 Registration and Operation of Enterprises

The places of operation of the enterprises registered in Nanjing and Ma'anhsan were analyzed separately. Kernel density analysis was also conducted on the places of registration of the enterprises operating in these two cities. It is discovered that 0.152% of the enterprises running in the two cities are registered in the other city while 0.688% of the registered enterprises are run in the other city. The enterprises in Nanjing and Ma'anhsan show insignificant connections between their places of registration and operation (Fig. 6).

4 Planning Applications Under Quantitative Analysis

4.1 Development Strategies for Integration

The relationships of Nanjing with Ma'anhsan and other cities in the metropolitan area were compared to discover the characteristics and advantages of integrated development of Nanjing and Ma'anhsan so that more effective integration development strategies could be formulated.

Under population integration, all types of connections can be enhanced. The connections between Ma'anhsan and Nanjing have become closer, and employment and business connections have gradually become stronger. Therefore, business and tourism connections can be strengthened to enhance the quality of the population integration of the two cities.

Under capital integration, more in-depth integration can be achieved progressively. The industrial structures of Nanjing and Ma'anshan are, to a certain degree, complementary to each other. Ma'anshan has received significant financing from Nanjing for manufacturing industry and scientific research/technological services. Hence, it is possible to build complementary industrial connections with Nanjing. More enterprises with great potential can be cultivated to promote industrial integration.

With the help of transportation integration, travelling has become more convenient. As the number of high-speed trains has increased, commuting efficiency by taking trains from Ma'anshan to Nanjing has improved. Similarly, it has become more convenient to commute between Nanjing and Ma'anshan. High-speed trains have promoted city-regional integration and reconstruction.

Lastly, it is worth exploring how to interpret the integration of the two cities based on the analysis of big data according to the Nanjing-Ma'anshan integration development framework. Further research is required to understand how to enhance human connections, change endogenous economic development, proactively establish two-way human connections and complementary industrial structures, and construct more convenient transportation systems.

4.2 Development Guidance for Key Areas

Based on the quantitative analysis framework, the key development blocks adjacent to Nanjing and Ma'anshan that are involved in their integration were extracted to calculate relevant indicators (Fig. 7). The results can provide more effective data support and development guidance for these areas.

City functions reflected by population integration and industrial characteristics reflected by enterprise properties were quantitatively analyzed. Six groups and 12 blocks involved in the integration of Nanjing and Ma'anshan were assessed for their characteristics to determine the foundation of integrated development. Based on the results and the overall situation in the region, development guidance for key areas was provided (Table 4).

5 Conclusions

Introducing big data into the regional integration research allows the identification of “invisible” elements and the analysis of their flow characteristics. As the sample size increases, more accurate integration research is possible. Big data cover a wide variety of dimensions as well as temporal and spatial ranges and can substantially increase the number of dimensions and the accuracy of the medium-scale and microscopic integration research.

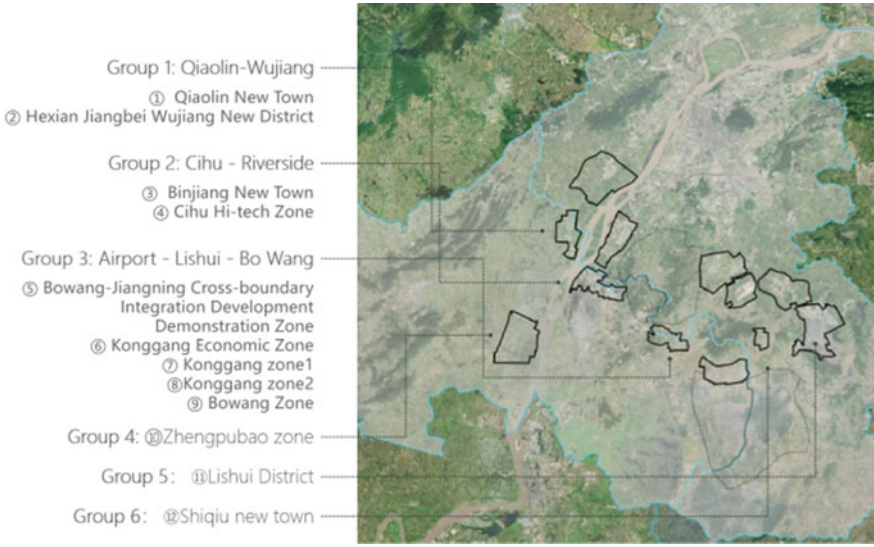


Fig. 7 Distribution map of Ning-Ma blocks

Table 4 Ning-Ma plate development evaluation table

Group number	District Grouping	City	Name of key areas	Evaluation of urban function attribute	Evaluation of industry development				Evaluation of group development	
					Industry size	Registered Business Environment	Dominant Industry	Degree of industrial concentration		Dominant Industry
1	Qiaolin-Wujiang	Nanjing	Qiaolin New District	Balance of employment and residence (mainly occupation)	Large scale	Strong registered environment	Scientific services	Higher concentration	Artificial intelligence integrated circuit	Qiao-Wu Occupation function integrated circuit Qiao - Wu functional complementarity and spill over
		Ma On Shan	Hexian Jiangbei Wu River New District	Residential attribute is dominant	Medium scale	Strong registered environment	Residential services	Weak agglomeration	Not obvious	
2	Cihu-Binjiang	Nanjing	Binjiang District	The dominant attribute employment	Large scale	Strong operating environment	Information services	Higher concentration	Artificial intelligence	Outstanding industrial function; Cihu has advantages in registered economy
		Ma'anshan	Cihu high-tech zone	Strong employment attribute	Medium scale	Strong registered environment	Manufacturing	Moderate concentration	Software information services	
3	Konggang-Lishui-Bowang	Nanjing/ Ma on shan	Bowang-Jiangning Demonstration area for cross-boundary integrated development	Balance of employment and residence (Employment-oriented)	Small scale	Balance of Notes (Good Business)	Manufacturing	Weak agglomeration	Not obvious	Nanjing has outstanding industrial function. Ma On Shan City is mature, "People for Land" basis
		Nanjing	Konggang economic zone	The strong attribute employment	Medium scale	Balance of Notes (Good Business)	Accommodation Catering	Weak agglomeration	Not obvious	
		Nanjing	Konggang district1	The strong attribute-Residential	Large scale	Strong operating environment	Manufacturing	Higher concentration	New energy vehicles Integrated circuit	
		Nanjing	Konggang district2	The strong attribute employment	Medium scale	Strong registered environment	Manufacturing	Moderate concentration	Biomedical medicine	
4	Zhengpubao New District	Ma on shan	Bowang District	The strong attribute-Residential	Large scale	Strong operating environment	Manufacturing	Weak agglomeration	Not obvious	Potential for Landmark Development
		Ma on shan	Zhengpubao New District	The strong attribute-Residential	Medium scale	Strong registered environment	Rental services, Scientific services	Moderate concentration	Software information services	
5	Lishui District	Nanjing	Lishui District	The strong attribute-Residential	Super scale	Strong operating environment	Wholesale and retail	Higher concentration	New energy vehicles, Integrated circuit	High productivity, Radiation-driven capability
6	Shiqiu new town	Nanjing	Shiqiu new town	The strong attribute employment	Medium scale	Balance of Notes (Good Business)	Recreation	Weak agglomeration	Not obvious	General level of development

Timely big data covering different dimensions can be used in regional integration research according to the research objectives. They can help reveal the integration characteristics more accurately to provide new ideas and support for integrated development planning.

References

1. Wang MY (2012) Revisiting and rethinking regional urbanization in Changjiang River Delta, China. *Chin Geogr Sci* 22(05):617–625
2. Li J, Chi S (2019) The impact of big cities of the new era on the economic growth of surrounding small and medium-sized cities (Chinese). *Sci Technol Progr Policy* 36(14):31–37
3. Yang C (2005) Multilevel governance in the cross-boundary region of Hong Kong–Pearl River Delta, China. *Environ Plan A* 37(12):2147–2168
4. Alderson AS, Beckfield J (2004) Power and position in the world city system. *Am J Sociol* (109):811–851
5. Hall P, Pain K (2006) *The polycentric metropolis: learning from mega-city regions in Europe*. Routledge Press, London, pp 91–125
6. Gao L, Jiang F (2017) Spatial reconstruction and core formation of regional integration development—a case study on Nanjiang-Zhengjiang-Yangzhou (Chinese). *Nanjing J Soc Sci* 07:41–48
7. Zhang A, Fan C (2019) Spatial autocorrelation and overall advancement of regional integration development in the Yangtze River Delta (Chinese). *Mod Econ Res* 08:15–24
8. Tang W (2011) Paths, models and countermeasures of urban-rural integration in West China (Chinese). Southwest University, Chongqing, pp 24–35
9. Xu J, Shi L (2010) Spatial dynamic mechanism and model framework of the integration development of the Hohhot-Baotou-Ordos Region (Chinese). *J Arid Land Resour Environ* 24(07):52–57
10. Fang C, Zhang Y (2014) Formation mechanisms, spatial organization models and patterns of urban integration areas in China (Chinese). *Urban Plan Forum* 06:5–12
11. Li P, Zhang X (2019) Research on the spatial structure characteristics and spatial integration development of the Yangtze River Delta (Chinese). *J Anhui Univ Technol (Soc Sci)* 02:148–156
12. Tang F, Wu P, Zhou L (2018) Measurement and evaluation of the degree of integration of Changsha-Zhuzhou-Xiangtan urban agglomeration (Chinese). *Econ Geogr* 38(02):59–65
13. Wu C, Huang X, Chen B et al (2017) Spatial patterns of economic ties and trends of economic integration in the Yangtze River Economic Belt (Chinese). *Econ Geogr* 37(07):71–78
14. Chen W, Wang J (2013) Preliminary measurement of the spatial integration development patterns of the Yangtze River Delta (Chinese). *Sci Geogr Sin* 33(08):902–908
15. Niu X, Li K (2019) Cross-city functional connections of the Shanghai Metropolitan Area from the perspective of a close one-day communication circle (Chinese). *Shanghai Urban Plan Rev* 03:16–22
16. Li X, Chen M (2016) Quantitative analysis method of integration under the background of big data and its application in Wenling City (Chinese). *Planner* 32(02):83–88
17. Pang Q, Jiang W, Luo Y, Man D, Chen J (2019) Integration and innovation big data technologies for urban master plan implementation and evaluation—taking Zhuhai as an example (Chinese). *Bull Surv Map* (06):155–159+162
18. Zhao Z, Wang B, Yang J (2020) Development, evolution characteristics and governance measures of spatial integration of the Yangtze River Delta based on satellite remote sensing big data (Chinese). *Dev Stud* 01:13–19

A Method of Household Car Ownership Prediction Using Ordered Probit Model



Guangzheng Yao, Yanyan Chen, Kaijun Cui, Donghui Xu, and Jiarui Liu

Abstract Predicting the demand for cars accurately is of great significance to urban planning and parking facilities planning. This paper uses the data from the Beijing 2010 Resident Travel Survey and takes the family income level, family demographic characteristics, the geographical environment of the family and the personal attributes of the head of the household as parameters. Random forest method is used to rank the importance of explanatory variables and carry out variable reorganization. Use the ordered probability model (Ordered Probit) to establish the relationship between explanatory variables and the number of vehicles owned by the household. The model found that the three parameters of the number of family driver's licenses, the level of family income, and the number of people have the greatest impact on the family's decision to own a car, and the income group mean transformation has a better fitting effect.

Keywords Urban transport · Household car ownership · Ordered Probit model · Random forest model

1 Preface

In recent years, with the continuous improvement of the income level of Chinese residents, more and more families can afford to buy cars. Private cars not only bring convenience to people's life, but also produce many negative problems. Firstly, due to the increase in the number of cars in Chinese cities and the mismatch between

G. Yao (✉) · Y. Chen

Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, No. 100 Pingleyuan, Chaoyang District, Beijing 100124, P. R. China

e-mail: yaogz@sohu.com

G. Yao · K. Cui · J. Liu

Nanjing Institute of City and Transport Planning Co. Ltd, Beijing, China

D. Xu

Beijing University of Civil Engineering and Architecture, Beijing, China

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

759

W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,

Lecture Notes in Electrical Engineering 775,

https://doi.org/10.1007/978-981-16-5429-9_57

the supply of urban parking facilities, there is a common phenomenon of random parking that affecting firefighting as well as damaging the living environment of the community and reducing the capacity and reliability of roads. Secondly, private car are highly related to using as a means of transportation. A large number of documents show that car ownership determines car use, and the use will cause traffic congestion, air pollution and energy consumption.

In China, there is no consensus on whether to restrict the purchase of cars or not. Some scholars call for proper control of the ownership and use of cars. They believe that private cars are personal items, car owners should bear the due internal and external costs for their ownership and use due to the negative externalities of its ownership, and the government should issue corresponding policies to minimize the ownership and use of cars. Many decision makers believe that owning a vehicle is a basic right of the citizens and the government should provide adequate transportation infrastructure including parking facilities. Different approaches will bring about different results. Under the conditions of different supply and restrictive conditions, the number of vehicles in cities is completely different. Which is a result of the demand for parking facilities differently as well as the land and funds requiring for the construction of parking facilities. Therefore, it is of great significance for urban planners and decision makers to predict the level of motor vehicle ownership and measure the corresponding investment reasonably.

Many foreign cities have established prediction models adapted to their respective cities based on the car ownership of local households, which lacks versatility and guidance for cities in China. However, domestic researches in China usually used the development trend of macro-historical data to predict the possible future demand. Due to the lack of judgment on the internal logical relationship between car development and urban management, it is difficult to guide policy practice. Therefore, it is necessary to establish a family-based car ownership prediction model suitable for Chinese cities.

2 Model

2.1 Research Review

Car ownership is a common concern of automobile manufacturers and urban planning managers. Developed countries such as Europe and the United States had experienced a period of rapid growth in the number of cars since the 1970s and 1980s. Early predictions of car ownership relied more on aggregate models, which mainly used linear regression, logistic function, the S-curve method and other methods to study the relationship between national and regional data and car ownership. The S-shaped growth curve usually has three phases: during the first phase, the car ownership increases slowly, and the per capita GDP is low at this stage; during the second phase, the car ownership rises rapidly, and the per capita GDP increases rapidly

during the third phase. The car ownership is close to the saturation level, and the per capita GDP growth at this stage tends to stabilize. A representative study of the S-shaped growth curve is that Dargay and Gately [1] selects per capita GDP as an independent variable and uses the Gompertz model to predict the car ownership rate per thousand people in 26 countries in 2015.

The disaggregate model is corresponding to the aggregate model. The disaggregate model used the family as the research unit to study the influence of different factors at the family and individual levels on the number of car ownership. Typical disaggregated models included sequential response models and non-sequential response models. The non-sequential response model believed that there is no strict sequential relationship between the number of family cars. Zegras [2], Caulfield [3] and other scholars used non-sequential response models to study car ownership. Potoglou and Kanaroglou [4] and others analyzed the data of the Hamilton area in Canada, taking the working status of the family population, family income, and the degree of land use around the place of residence as influencing factors. Savolainen et al. [5] believed that compared with the sequential response model, the non-sequential response model had higher flexibility in model construction. On the contrary, the sequential response model believed that there is a progressive relationship between the number of cars in the family. Scholars such as Potoglou and Susilo [6] and Ko [7] believed that the sequential response model has better explanatory properties. Kim [8] used a sequential response model to analyze the impact of the number of driver's licenses, family size, and residential area on the number of cars. Chu [9] took the family income, the number of family population, the population density around the place of residence, and the degree of land use as Influencing factors, Huang et al. [10] took family income, the number of driver's licenses, public transportation and other factors as influencing factors, and focused on the analysis of the impact of urban rail transit on the number of family cars.

In addition to the sequential and non-sequential response models, Pinjari [11] used a multi-discrete continuous generalized extreme value model to analyze the impact of family population work and income on family car ownership. Clark [12] used a geographically weighted regression model to analyze the impact of population density and resident income in different parts of the UK on car ownership. Clark [13] used the naive Bayes algorithm in the field of data mining to analyze the impact of family characteristics and family member characteristics on car ownership. Yang et al. [14] used a two-level planning method to explore the relationship between car ownership and road carrying capacity, regional traffic demand and traffic flow. Yagi and Managi [15] used dynamic regression methods to analyze the demographic characteristics of Japan and the impact of demographic changes on car ownership and demographic characteristics.

2.2 Ordered Probit Model

The Ordered Probit model assumed that the observed variables that affected the decision-making variables can be constructed as a linear function, and the value of the constructed function reflects the willingness of households to purchase vehicles. The more larger value, the higher the willingness to buy a vehicle.

The Ordered Probit model specific function form is as follows:

$$Y_i^* = \sum_{k=1}^K \beta_k X_{ik} + \varepsilon_i \quad (1)$$

Y_i^* represents the potential willingness of the family to own a vehicle, X represents the explanatory variable vector, β represents the parameter vector to be estimated, ε represents random items.

The value of Y_i^* in each family represents the potential willingness of the family to own a vehicle, but it is difficult to directly observe. However, the number of vehicles in a family can be observed, and an ordered probability model is established by establishing the relationship between Y_i and Y_i^* . For example, according to the level of car ownership, households can be divided into “households without a car” and “households with a car”, and “households with a car” can be subdivided into households with 1 car, 2 cars, 3 cars, etc. The variable of Y_i is the amount of car ownership in the family, $Y_i = 0$ which means that the family does not have a car, $Y_i = 1$ means that the family has 1 car, $Y_i = 2$ means that the family owns 2 cars, and $Y_i = 3$ means that the family owns more than 3 cars. The potential willingness of family to own private cars has the following correspondence with the level of family car ownership:

$$\begin{aligned} Y_i = 0, & \quad \text{if } Y_i^* \leq \delta_1 \\ Y_i = 1, & \quad \text{if } \delta_1 \leq Y_i^* \leq \delta_2 \\ Y_i = 2, & \quad \text{if } \delta_2 \leq Y_i^* \leq \delta_3 \\ Y_i = 3, & \quad \text{if } Y_i^* \geq \delta_4 \end{aligned} \quad (2)$$

δ is the intercept term, which means that the classification of family car ownership depends on whether the car ownership tendency crosses a certain threshold. The probabilities of car ownership tendency for the values 1, 2, and 3 are expressed as follows:

$$\begin{aligned} \Pr(Y_i = 0) &= \Pr(Y_i^* \leq \delta_0) \\ \Pr(Y_i = 1) &= \Pr(\delta_0 \leq Y_i^* \leq \delta_1) \\ \Pr(Y_i = 2) &= \Pr(\delta_1 \leq Y_i^* \leq \delta_2) \\ \Pr(Y_i = 3) &= \Pr(Y_i^* \geq \delta_2) \end{aligned} \quad (3)$$

Table 1 Number and classification of family cars

The number of car ownership	Frequency	Percentage	Cumulative percentage
0	33,292	70.99	70.99
1	12,571	26.80	97.79
2	986	2.10	99.89
3	51	0.11	100
Total	46,900	100	

3 Data

3.1 Data Description

This paper adopted the data of the fourth resident travel survey conducted in Beijing in 2010, including 265 neighborhoods and towns in 16 districts and counties covering the whole city. The survey was sampled at a ratio of two-thousandths of the total population of the city, and the samples were evenly distributed in each area. The method of stratified random sampling was adopted to collect data from more than 40,000 households. The survey content included family background information, basic information of family members, family members travel and travel location table, residents' daily travel log and so on. In the original data of the resident travel survey, the family background information table included the family number, community number, number of cars owned by the family, income status, permanent residents, number of workers, housing type, number of non-motor vehicles, etc. The content of the family member table included the family number, member number, gender, household registration type, age, occupation, number of driver's licenses, educational background, and relationship with the head of a family member.

The number of cars owned by households was used as a decision variable, and the distribution was shown in Table 1. Car-free households accounted for about 71%, households who bought one car accounted for about 26%, households who bought two cars accounted for about 2.10%, and households who bought three cars or more accounted for about 0.11%.

3.2 Parameter Description

From the questionnaire, the table of family background information and the family member basic information were integrated, and a total of 16 factors that have an impact on car ownership were extracted, including family income, residential area, family size, etc. Variable statistics such as Table 2 shows.

Table 2 Statistical table of explanatory variable

Variables	Min	Max	Mean	Std. dev.	Remarks
PeoplNum	1	5	2.471642	0.9758735	
WorkNum	0	4	1.178443	0.91831	
FullJob	0	4	1.136482	0.9176802	
PartJob	0	3	0.0390618	0.2105168	
Kid_Not	0	1	–	–	Categorical variable
DistrictType	1	4	–	–	Categorical variables
Bike	0	5	0.9639863	0.9288447	
Ebike	0	4	0.1542029	0.401609	
Motor	0	3	0.0309318	0.1822833	
IncomeLevel	1	8	–	–	Ordinal variable
Gender	1	2	–	–	Categorical variables
Age (dis)	1	10	–	–	Ordinal variable
HujiType (order)	0	1	–	–	Categorical variables
Zuigaoxueli (order)	0	8	–	–	Ordinal variable
RenyuanType (order)	1	3	–	–	Categorical variables
LicenceNum	0	5	0.7948614	0.849132	

- (1) Family income level: The family income level was the income range of the total annual income of the family members as a unit, and was an ordinal categorical variable with a value of 1–8. The corresponding income ranges are: 0–50,000, 50,000–100,000, 100,000–150,000, 150,000–200,000, 200,000–250,000, 250,000–300,000, 300,000–500,000 and more than 500,000. Since the ordinal variables cannot reflect the household income well, this assay will compare the effect of the income ordinal categorical variables, the expectations of the income grouping interval, and the model effect after logarithmization.
- (2) The urban area where the family located: In order to analyze the impact of location on the family’s car buying behavior, the urban area variable of the family is set as an ordered variable according to the spatial position relationship between the county and the central urban area, which is divided into four categories: Dongcheng District and Xicheng District represented the value of is 1; the value 2 included Chaoyang District, Haidian District, Fengtai District and Shijingshan District; the value of Shunyi District, Tongzhou District, Daxing District, Fangshan District, Mentougou District and Changping District is 3; Yanqing District, Pinggu District, Huairou District and Miyun District take a value of 4.
- (3) Children: Whether there were children in the family was set as categorical variables, by taking the period of survey as the node to count whether there are children younger than 6-year-old in the family, “Yes” is 1 and “No” is 0.

- (4) Gender of the head of household: This variable was set as a categorical variable, 1 indicated that the gender of the head of household is male, and 2 indicated that the gender of the head of household is female.
- (5) Age grouping of household heads: This variable was set as ordinal variable, by grouping at 10-year-old intervals.
- (6) Household registration type of the head of household: This variable was set as categorical variables, 1 was local household registration, 0 was non-local household registration.
- (7) The highest educational background of the head of household: This variable was set as an ordinal variable, 0 meant no education, 1–8 are preschool children, elementary school, junior high school, high school, technical secondary school, college, undergraduate, graduate and above.
- (8) Working status of the head of household: This variable was set as categorical variables, 1 ~ 3 indicated that the working status of the head of household is full-time, part-time and non-working status respectively.

4 Parameter Estimation

4.1 Parameter Selection

Due to the large number of explanatory variables, it is necessary to simplify the model and extract the key influential variables in the actual prediction process. In this paper, the Random Forest Method was used to screen parameters. The Random Forest Method is a kind of combination classifier, its mechanism is to use the appropriate sampling method to extract multiple samples from the original samples to construct decision tree model, and then these decision tree models are combined to obtain the final classification results by using method of vote.

The number of decision trees built in Random Forest is usually determined by error graph (Fig. 1). In this paper, the number of decision trees was 150. After substituting all the explanatory variables into the Random Forest, the importance of each parameter was calculated (Fig. 2).

In order to avoid multicollinearity, the correlation between variables can be analyzed by heat map (Fig. 3). By estimating the Pearson correlation coefficient between the variables and the family car ownership. It can be found that the number of family driving licenses is most closely related to the number of family car ownership. It can be found that variables such as family income level, the number of family members and the number of family workers are positively correlated with the number of family car ownership, while the number of non-motor vehicles (bicycles, electric vehicles, motorcycles), the gender and age of householder are negatively correlated with the number of family car ownership.

It is worth noting that the explanatory variables such as family workers number, the number of full-time workers and part-time workers are highly correlated, which is the same as bicycles number, electric vehicles number and motorcycles number.

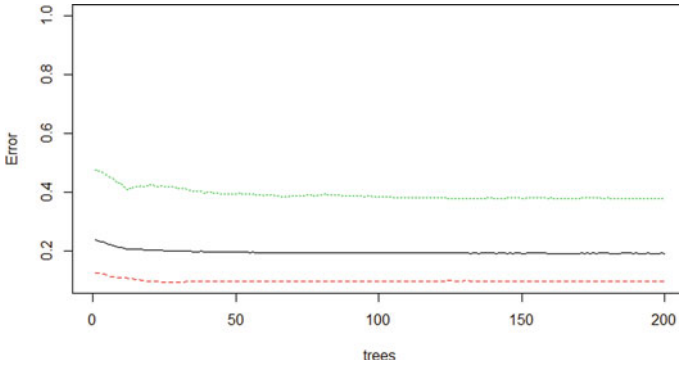
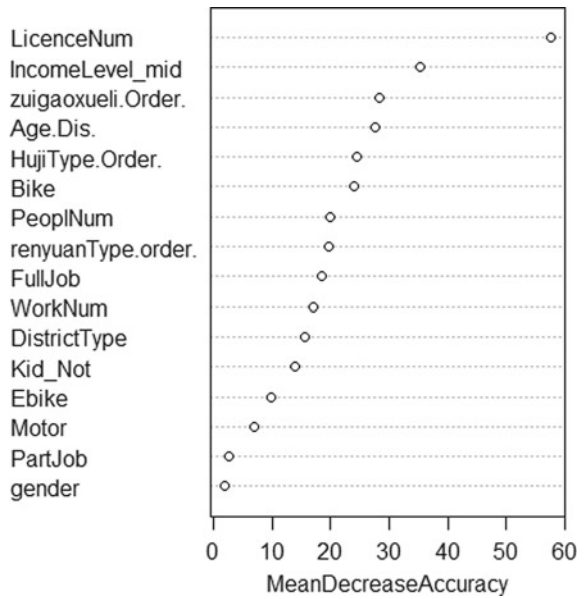


Fig. 1 Error graph of random forest decision tree

Fig. 2 The importance map of random forest parameter



It can be combined with the importance of variables for screening, and suitable variables can be selected to be substituted into the model for fitting. By sorting the importance of the variables from low to high, the appropriate variables are selected and substituted into the model for fitting.

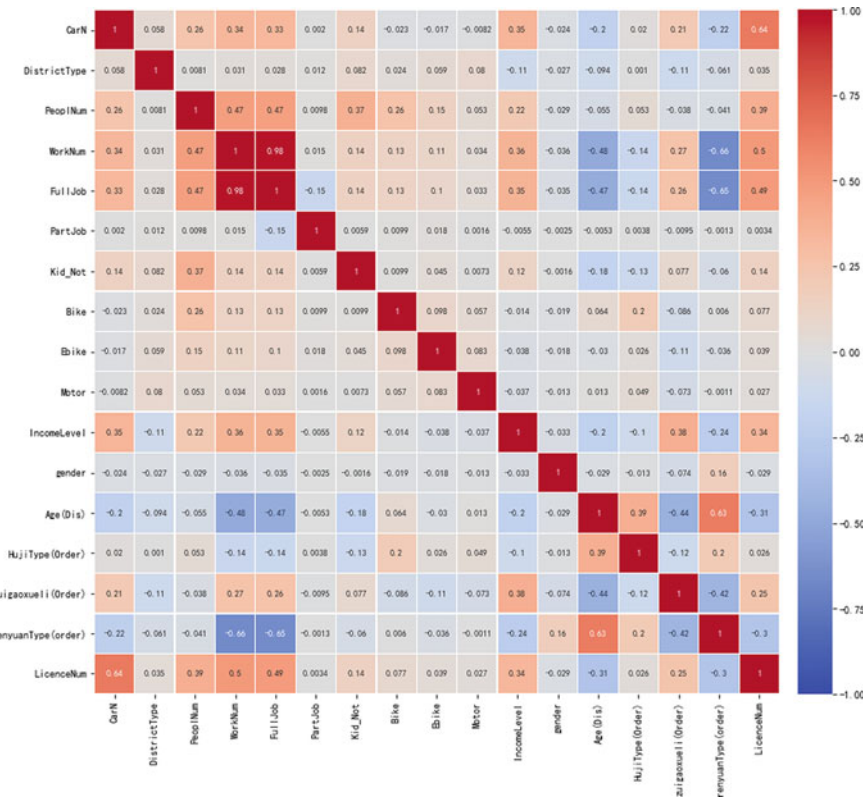


Fig. 3 Heat map of explanatory variables and family car ownership

4.2 Model Construction and Parameter Calibration

This article compares the effects of the four models. The original data is substituted into the Ordered Probit model, and the parameters are calibrated by the maximum likelihood method (Model 1). With comparing the different fitting degree of different models by using the method of random forest and heat map to filter the variables (Model 2) as well as different treatment of income variables such as replacing the income variable with its expectations (Model 3) and algorithmized the income variable (Model 4), the essay expected to get the best model to predict the car ownership in different families (Table 3).

The estimation of models 1 showed that there were four explanatory variables had no significant impact among 16 explanatory variables such as the number of full-time staff in family, the number of part-time staff in family, the gender and the age of the head of the household. Considering the importance of the variables obtained method by the random forest and correlation between variables by the heat map, Model 2 is obtained after filtering explanatory variables. The explanatory variables of Model 2

Table 3 Model parameter calibration results

Variable	Model 1		Model 2		Model 3		Model 4	
	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value
DistrictType	0.1074	11.6809	0.1031	11.3706	0.1193	13.0712	0.1348	14.6419
PeoplNum	0.063	6.1219	0.056	5.523	0.0512	5.0276	0.0433	4.2326
WorkNum	0.0502	2.5251	0.0359	2.8138	0.0143	1.1139	-0.0146	-1.1278
FullJob	-0.0301	-0.3213	-	-	-	-	-	-
PartJob	0.0066	0.079	-	-	-	-	-	-
Kid_Not	0.1594	6.3922	0.1569	6.3866	0.1480	5.9909	0.1433	5.7744
Bike	-0.1538	-18.7942	-0.1581	-19.396	-0.1523	-18.6278	-0.1510	-18.4087
Ebike	-0.1533	-8.4339	-	-	-	-	-	-
Motor	-0.1597	-4.1588	-	-	-	-	-	-
IncomeLevel_mid	0.0326	11.6113	0.0326	11.6613	0.0400	25.1350	0.5864	32.4811
Gender	0.0011	0.0713	-	-	-	-	-	-
Age (dis)	0.0015	0.2239	-	-	-	-	-	-
HujiType (order)	0.1247	6.0758	0.1165	5.97	0.1481	7.5073	0.1723	8.6691
Zuigaouxeli (order)	0.0432	9.0176	0.0481	10.44	0.0286	6.0713	0.0071	1.4648
RenyuanType (order)	-0.0526	-4.5642	-0.0505	-4.8399	-0.0546	-5.2096	-0.0657	-6.2319
LicenceeNum	1.1311	103.5511	1.1287	104.6038	1.1171	102.8489	1.1072	101.4207
Cons	-2.2475	-36.983	-2.2498	-43.7321	-2.3545	-45.3756	-2.9724	-52.1755
Cut1	0.7896	90.9681	0.786	90.8372	0.8078	92.4517	0.8196	93.7544
Cut2	0.3723	9.8919	0.3563	9.464	0.4134	10.8963	0.4034	10.7318
Rectified Mac-ratio	0.8413		0.8051		0.8075		0.8095	

are all significant, which showed that this model was more explanatory than Model 1. The driver’s license had the greatest impact on family car ownership which it’s t-value was 104.6038. The possible reason was that the license determines whether to drive or not. The estimation of income is positive which meant the higher income the more likely to purchase cars. The larger the family population and the greater the number of working people, the greater the number of car ownership in the family. The possible reason was that these factors cause the greater the travel demand and the higher the willingness to purchase cars.

In addition, Model 3 replaced the income variable with the expectation of the income interval to obtain based on Model 2. Model 4 algorithmized the income variable based on Model 3. By comparing the fit degree of the model (Rectified Mac-ratio), it can be found that the fit degree of the expected model after the income interval is increased by 0.0024 compared with the raw data. After the expected legitimization of the income interval, the fit of the model is increased by 0.0045 compared to before the variable of income was not processed. It was shown that the legitimization of the income variable can better illustrate the impact of income on the number of car ownership. Although the fit of model 4 is slightly lower than that of model 1, the fit degree was still above 80% which indicating that the model works well.

4.3 Model Evaluation

In order to verify the effectiveness of the model after variable selection and combination, Table 4 showed the prediction results. The fitting degree of model (Rectified Mac-ratio) was 80.95%. The accuracy of the model for unbalanced data sets was usually measured by indicators such as Precision, Recall, and Micro-average_F1. This article selects Micro-average_F1 as the measure of model accuracy. The accuracy of the model to predict 0 vehicles is 87.29%. The absolute value of the maximum likelihood estimate of the model is 20,515.097, the model Prob > $\chi^2 = 0.0000$ and all explanatory variables pass the significance test, indicating that the model passes the significance test.

Table 4 Model prediction results

Prediction	Reality				Total of prediction	Prediction accuracy (%)
	0	1	2	3		
0	29,288	4477	41	1	33,807	87.29
1	3881	7517	745	19	12,162	60.79
2	119	559	183	26	887	19.54
3	4	18	17	5	44	10.53
Total of reality	33,292	12,571	986	51	46,900	78.88

Table 5 Analysis of marginal utility of explanatory variables

	Marginal utility	Std.	z	$P > z $
LicenceNum	-0.23874	0.001223	-195.2	0
IncomeLevel ~ d	-0.0097	0.000331	-29.28	0
Non_motor_v ~ e	0.053842	0.00306	17.59	0

The marginal utility analysis of the variables after the combination screening is shown in Table 5. Compared with other variables, the marginal utility of variables such as the number of driving licenses, household income level and the number of non-motor vehicles is significant. When the number of family driving licenses increases by 1, the probability that the family car ownership is still zero will drop by 23.8%. When the family income increases by 10,000 yuan, the probability that the family car ownership is still zero will drop by 0.9%.

5 Summary

Exploring the influencing factors of family car ownership and predicting the number of car ownership accurately is of great significance to urban transportation planning and policy making. Due to the sequential relationship between the number of family cars, the sequential response model was used to analyze the number of family cars, and 16 explanatory variables were divided into four categories: family income level, family demographic characteristics, and family location Environment and personal attributes of the head of the household. Using random forest to calculate the importance of each explanatory variable, the study found that there are 12 parameters that have a significant impact on the number of family cars. Among them, the number of driving licenses, family income, family population, and the number of employees have positive effects, indicating that the number of car ownership is the same as their changing trend. While the impact of the number of non-motorized vehicles such as bicycles is negative, indicating that the increase in the number of non-motorized vehicles will weaken the demand for households to buy private cars.

Through comparison with different models, this essay found that the legitimization of the income interval can better explain the impact of income on household car ownership. Through random forest algorithm and heat map analysis, combined with the output of the Ordered Probit model, it is considered that 10 variables including the family district type, the work stuff, whether the family has children under 6-year-old, the number of bicycles, the level of family income, HujiType, highest degree and working status of household head and the number of driver's licenses, which can well predict the number of car ownership in the family.

Acknowledgements Fund Projects: Project funded by Beijing Natural Science Foundation (9174027).

References

1. Dargay J, Gately D (1999) Income's effect on car and vehicle ownership, worldwide: 1960–2015. *Transp Res Part A Policy Pract* 33(2):101–138
2. Zegras C (2010) The built environment and motor vehicle ownership and use: evidence from Santiago de Chile. *Urban Stud* 47(8):1793–1817
3. Caulfield B (2012) An examination of the factors that impact upon multiple vehicle ownership: the case of Dublin, Ireland. *Transp Policy* 19(1):132–138
4. Potoglou D, Kanaroglou PS (2008) Modelling car ownership in urban areas: a case study of Hamilton, Canada. *J Transp Geogr* 16(1):42–54. <https://doi.org/10.1016/j.jtrangeo.2007.01.006>
5. Savolainen PT, Mannering FL, Lord D et al (2011) The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. *Accid Anal Prev* 43(5):1666–1676
6. Potoglou D, Susilo Y (2008) Comparison of vehicle-ownership models. *Transp Res Rec J Transp Res Board* 2076:97–105
7. Kim D, Ko J, Park Y (2015) Factors affecting electric vehicle sharing program participants' attitudes about car ownership and program participation. *Transp Res Part D* 36:96–106. <https://doi.org/10.1016/j.trd.2015.02.009>
8. Hong SK, Kim E (2010) Effects of public transit on automobile ownership and use in households of the USA. *Rev Urban Reg Dev Stud* 16(3):245–262
9. Chu Y-L (2002) Automobile ownership analysis using ordered probit models. *Transp Res Rec J Transp Res Board* 1805:60–67
10. Huang X, Cao X (Jason), Yin J et al (2017) Effects of metro transit on the ownership of mobility instruments in Xi'an, China. *Transp Res Part D Transp Environ* 52:495–505. <https://doi.org/10.1016/j.trd.2016.09.014>
11. Pinjari AR (2011) Generalized extreme value (GEV)-based error structures for multiple discrete-continuous choice models. *Transp Res Part B Methodol* 45(3):474–489. <https://doi.org/10.1016/j.trb.2010.09.004>
12. Clark SD (2007) Estimating local car ownership models. *J Transp Geogr* 15(3):184–197. <https://doi.org/10.1016/j.jtrangeo.2006.02.014>
13. Clark SD (2009) The determinants of car ownership in England and Wales from anonymous 2001 census data. *Transp Res Part C Emerg Technol* 17(5):526–540. <https://doi.org/10.1016/j.trc.2009.04.004>
14. Yang Z, Chen G, Yu B (2008) Car ownership level for a sustainability urban environment. *Transp Res Part D Transp Environ* 13(1):10–18. <https://doi.org/10.1016/j.trd.2007.10.003>
15. Yagi M, Managi S (2016) Demographic determinants of car ownership in Japan. *Transp Policy* 50:37–53

Long Short-Term Memory Neural Network for Traffic Speed Prediction of Urban Expressways Using Floating Car Data



Deqi Chen, Xuedong Yan, Shurong Li, Xiaobing Liu, and Liwei Wang

Abstract Recently, accurate and rapid prediction of traffic speed has become a research hotspot in the current applied traffic field. Based on deep learning models, especially Long Short-Term Memory neural network (LSTM), real-time traffic speed prediction of the urban expressway has been a new challenge. Since it takes a long time to extract traffic parameter by map matching method, we present a novel grid model to rapidly derive a series of traffic parameters from the floating car data (FCD). To improve the prediction accuracy, we consider the spatial–temporal characteristics of traffic speed (i.e. upstream and downstream speed, historical average speed, etc.). To verify the validity of the model of the proposed model, we employ 40-day FCD to train and test the model. Our final result, compared with other machine learning methods, the LSTM model has advantages of accuracy and stability, which could facilitate the prediction of the traffic speed and the traffic operation performance.

Keywords Long short-term neural network · Grid model · Traffic speed prediction · Floating car data

D. Chen · X. Yan (✉) · S. Li · X. Liu · L. Wang
MOT Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, P. R. China
e-mail: xdyan@bjtu.edu.cn

D. Chen
e-mail: 17114230@bjtu.edu.cn

S. Li
e-mail: 18114040@bjtu.edu.cn

X. Liu
e-mail: 16114222@bjtu.edu.cn

L. Wang
e-mail: wangliwei@bjtu.edu.cn

1 Introduction

Recently, traffic speed prediction has been one of the critical needs of the Intelligent Transport System. Through the accurate short-term traffic speed prediction, it can not only provide timely traffic information for transportation managers but also can optimize the route schedule of drivers [18].

With the rapid development of big data technology, the real-time traffic data (such as loop-detected-data, video data, microwave data and so on) for detecting traffic speed can be obtained by fixed and mobile sensors [9, 15, 30]. Above all, the Floating Car Data (FCD) based on mobile sensors has become a reliable, economical, and effective method [31]. Floating vehicles will regularly send the position and speed information to the data centers, where cleans and transforms the trajectory points [10]. Compared with the traditional data collection method, the FCD has the advantages of high quality, low cost, and wide coverage [17], which has become a new frontier [22]. Therefore, we can extract the traffic speed parameters from FCD to prepare for prediction. In traffic parameter extraction, map matching is a basic method. However, it manifests two limitations: time-consuming and lack of high-precision of the map. To overcome these limits, a novel grid model is proposed to extract the traffic speed parameter [5]. Later, the traffic parameters (i.e. upstream and downstream speed, history average speed, and standard deviation of speed, etc.) are constructed the time series of speed as the input dataset.

In recent decades, short-term traffic forecasting has been attracted by the extensive attention of researchers all over the world. It can be divided into two categories: the parametric prediction and the non-parametric prediction [24]. The most representative examples of nonparametric prediction methods are Support Vector Regression (SVR) [6] and K-Nearest Neighbor (KNN) [3]. KNN model has been widely used in the traffic flow prediction, which many studies have successfully applied to future traffic forecasting [28]. However, the size and integrity of the case database have a great influence on the prediction performance of the KNN model. SVR is another widely used method in traffic forecasting. Although Vanajakshi and Rilett found that the performance of SVR was better than that of neural network in the case of the small training dataset [23], it is limited by the lack of sufficient traffic patterns, which leads to poor long-term prediction accuracy [12, 23]. In order to capture the dynamic characteristics of traffic speed, Recurrent Neural Network (RNN) is applied to the spatial-temporal evolution of traffic speed. However, there are two shortcomings related to the traditional RNN model: (1) most RNN must be pre-determined in time step; (2) traditional RNN cannot deal with the long-term input sequence [18]. In order to overcome these defects, the RNN structure variant Long Short-Term Memory neural network (i.e. LSTM) was proposed by Hochreiter and Schmidhuber [14], which can be better adapted to the long-short term traffic speed prediction. By comparison with the RNN, LSTM has even broader applications: image captioning [21], speech recognition [32] and wind power forecasting [29], etc. Meanwhile, in the field of traffic prediction, LSTM is gradually becoming a popular research topic,

such as traffic flow prediction [11, 27], short-term Traffic Prediction [1], traffic speed prediction [18].

Compared with the existing literature, the proposed method has the following major advantages: (1) Under the grid modeling, the roads are transformed into discrete grids, a series of key traffic parameters (i.e. upstream and downstream speed, history speed, the standard deviation of speed, average speed and so on) can be rapidly extracted without complex algorithm. (2) The LSTM with the grid model can sufficiently capture the spatial-temporal characteristics of traffic speed for the real-time speed prediction of different time steps. (3) This method is easily transferred to other cities for the prediction of traffic speed.

The rest of the paper is organized as follows. Section 2 describes the methodology including data cleaning, construction of time series, and LSTM modeling. Section 3 takes Beijing Expressway as an example, and the experimental results are analyzed in detail. Finally, Sect. 4 presents the major conclusions and future research.

2 Methodology

2.1 Data Collection and Cleaning

Through the data center, we obtained a large number of FCD, whose rate is 1 min. As you can see, the base attributes of FCD are shown in Table 1, mainly including five characteristics: Terminal ID (*id*), GPS timestamp (*t*), Longitude (*lng*), Latitude (*lat*), ground speed (*v*).

During the FCD collection and the transmission process, there may contain some systematic mistakes, caused by GPS signal blockage or hardware/software errors [8]. Therefore, the methods of summarizing and filtering the collected FCD are as follows:

- (1) Redundant data: there are extremely similar data or duplicate data in the acquired FCD. Generally, there are multiple sampling data on the same time node, and redundant sampling data needs to be eliminated.
- (2) Coordinate longitude and latitude out of range: remove the FCD which is out of range of traffic road.

Table 1 The FCD attributes

Characteristic	Field name	Field type	Field description
<i>id</i>	Terminal ID	String	6 bytes characters, marking each vehicle
<i>t</i>	GPS timestamp	Timestamp	Unit: second
<i>lng</i>	Longitude	Floating	Unit: angle
<i>lat</i>	Latitude	Floating	Unit: angle
<i>v</i>	Ground speed	Integer	Kilometer per hour

- (3) Speed exceeds the threshold: according to the highest setting of the urban road traffic speed and the actual situation, we should eliminate the FCD which is beyond the range of 0–120 km/h.
- (4) Data missing: use the exponential smoothing method to supplement the missing data.

2.2 Construction of Time Series

Based on the grid model, a road network is constructed to transform the expressway into discrete cells, and the FCD matched to the grids through a simple assignment process [13]. This method avoids the complex and time-consuming calculation of map-matching algorithms as well as the needs of a detailed and timely updated map, leading to a considerable improvement in the efficiency of traffic operation evaluation. Considering the road construction, we empirically estimate the size of the grid $100 \times 100 \text{ m}^2$, reconstructing the urban expressway network [5].

Traffic speed prediction at one location normally uses a sequence of speed values with n historical time steps as the input data, which can be extracted by FCD. Considering that the traffic speed in grid g may be influenced by the velocities of nearby grids or even grids far away, especially when traffic jam propagates through the traffic grid-network [7]. To take these spatial influences into account, the proposed grid models adopt the upstream and downstream traffic speed. Besides, when predicting the grid-wide traffic speed, the impact of history average speeds and median speed on each location in the traffic network should not be neglected. So we selected six parameters: upstream average speed (i.e. U_t^g), downstream average speed (i.e. D_t^g), the standard deviation of traffic speed (i.e. S_t^g), historical median speed (i.e. M_t^g), historical average speed (i.e. H_t^g), average speed (i.e. V_t^g) of in time step t (see Table 2) to predict the average speed in future time step (i.e. $t + 1, t + 2, t + 3$), respectively. According to the grid-wide network, we extracted the six variables (i.e. $U_t^g, D_t^g, S_t^g, M_t^g, H_t^g, V_t^g$) of multivariate time intervals, respectively.

Table 2 The parameter definition

Parameter notation	Definition
U_t^g	Upstream average speed of time step t in grid g
D_t^g	Downstream average speed of time step t in grid g
S_t^g	Standard deviation of speed of time step t in grid g
M_t^g	Historical median speed of time step t in grid g
H_t^g	Historical average speed of time step t in grid g
V_t^g	Average speed of in time step t in grid g

Supposed that the traffic network consists of grids, and the input dataset is composed of six variables ($U_t^g, D_t^g, S_t^g, M_t^g, H_t^g, V_t^g$). The input dataset can be characterized as a data matrix (i.e. X_T^g), and the output can be expressed as a vector (i.e. Y_T^g). The time series of dataset matrix $[X_T^g, Y_T^g]$ as follows:

$$[X_T^g, Y_T^g] = \begin{bmatrix} U_{t-n}^g, & D_{t-n}^g, & S_{t-n}^g, & M_{t-n}^g, & H_{t-n}^g, & V_{t-n}^g, & \widehat{V}_{t-n+1}^g \\ U_{t-n+1}^g, & D_{t-n+1}^g, & S_{t-n+1}^g, & M_{t-n+1}^g, & H_{t-n+1}^g, & V_{t-n+1}^g, & \widehat{V}_{t-n+2}^g \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ U_{t-1}^g, & D_{t-1}^g, & S_{t-1}^g, & M_{t-1}^g, & H_{t-1}^g, & V_{t-1}^g, & \widehat{V}_t^g \\ U_t^g, & D_t^g, & S_t^g, & M_t^g, & H_t^g, & V_t^g, & \widehat{V}_{t+1}^g \end{bmatrix} \quad (1)$$

So, based on the data matrix $[X_T^g, Y_T^g]$, it is further split into the training dataset (i.e. $[X_T^g, Y_T^g]_{tr}$, whose proportion is 70%) for training and the test dataset (i.e. $[X_T^g, Y_T^g]_{te}$, whose proportion is 30%) for testing.

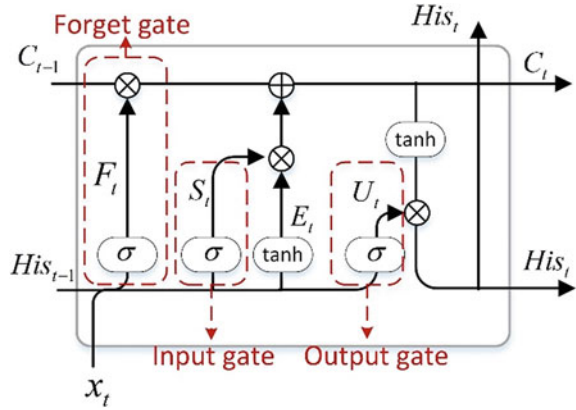
To improve the effectiveness of the optimal solution by eliminating, the data normalization is proposed, which using the proportion of calculation makes the dataset fall into the specific range and removes the limitation of data units, to convert the dataset to no dimensional pure data for the model training and testing [26]. Additionally, under excluding the effects of large eigenvalues, normalization can improve the training speed. The minimum–maximum method (see formula 2) is one of the major normalized methods.

$$X'_T = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (2)$$

2.3 LSTM Modelling

Being similar to other deep learning structures, LSTM has two stages: the training stage and the prediction stage. During the training stage, according to the hidden layers capturing and release the memory of historical traffic speed characteristics, LSTM can make a decision whether storing historical information or not and calculate the weights, biases of traffic parameters. From Fig. 1, it can be seen that the LSTM consists of three gates: input gate, forget gate, and output gate, which can determine whether the historical traffic information is transmitted to the next node. Meanwhile, it also can improve the efficiency of mining historical speed features and reduce the training time. During the prediction stage, based on weight and biases from the training stage, with the method of vector calculation of input dataset, we obtain the speed prediction over various time steps.

Fig. 1 LSTM framework



The LSTM transition equations are the following:

Gates:

$$F_t = \sigma(W_f \cdot [H_i s_{t-1}, x_t] + B_f) \tag{3}$$

$$S_t = \sigma(W_s \cdot [H_i s_{t-1}, x_t] + B_s) \tag{4}$$

$$E_t = \sigma(W_e \cdot [H_i s_{t-1}, x_t] + B_e) \tag{5}$$

Input transform:

$$C_t = F_t \cdot C_{t-1} + E_t \cdot S_t \tag{6}$$

Memory update:

$$U_t = \sigma(W_U \cdot [H_i s_{t-1}, x_t] + B_U) \tag{7}$$

$$H_i s_t = U_t \cdot \tanh(C_t) \tag{8}$$

Activation functions:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{9}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{10}$$

where x_t denotes the input vector at the time step t ; W_f, W_s, W_e present the weights; B_f, B_s, B_e present the biases; $\sigma(x)$ is the Sigmoid function range from 0 to 1; $\tanh(x)$ is non-linear activation functions [16].

During the training stage, the Rectified Linear Unit (*ReLU*) is proposed to better solve the problem of exploding/vanishing gradient and make converge fast [19].

$$ReLU = \begin{cases} x, & x > 0 \\ 0, & x < 0 \end{cases} \tag{11}$$

2.4 Index of Performance

To compare the effectiveness of the LSTM with the benchmark model, four criteria are proposed: the Mean Absolute Error (MAE), the Root-Mean-Square Error (RMSE), the Mean-Square Error (MSE) and R-square (R^2). The definitions are as follows:

$$MAE(Y_i, \hat{Y}_i) = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \tag{12}$$

$$RMSE(Y_i, \hat{Y}_i) = \left[\frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|^2 \right]^{\frac{1}{2}} \tag{13}$$

$$MSE(Y_i, \hat{Y}_i) = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|^2 \tag{14}$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y}_i)^2} \tag{15}$$

where Y_i is the real traffic speed; \hat{Y}_i is the traffic speed to be predicted; n is the number of test samples.

3 Result Analysis and Compassion

3.1 Data Description and Experimental Setup

To verify the proposed model, we select a section of the Western 3rd Ring-road expressway as the prediction research object (see Fig. 2). This road segment consists of 23 grids (i.e. the length is 2.3 km). The traffic speed in this area changes obviously in the daytime, which is more conducive to test the stability of the model.

With the grid model, we derive six key traffic parameters ($U_i^g, D_i^g, S_i^g, M_i^g, H_i^g, V_i^g$). The floating cars passing through this section are divided into the southbound and northbound directions. Later, set 10 min as the time slice, which means 144-time slices a day. Therefore, the dimension matrix of $[X_T^g, Y_T^g]$ is $[144 \times 40, 7]$. According to the rule of the dataset classification, the training dataset (i.e. $[X_T^g, Y_T^g]_{tr}$) is $[144 \times 40 \times 70\%, 7]$, and the test dataset (i.e. $[X_T^g, Y_T^g]_{te}$) is $[144 \times 40 \times 30\%, 7]$. According to the formula minimum–maximum normalization method, we normalized the training dataset, the test dataset. In this LSTM network, the dimension of the hidden unit is set to 150 and the dimension of mini-batches is supposed to 64, and the epochs are set to 150. For the activation function, the Rectified Linear Unit (i.e. *ReLU*) is proposed. Meanwhile, the Adam algorithm is used to update the parameters of LSTM.

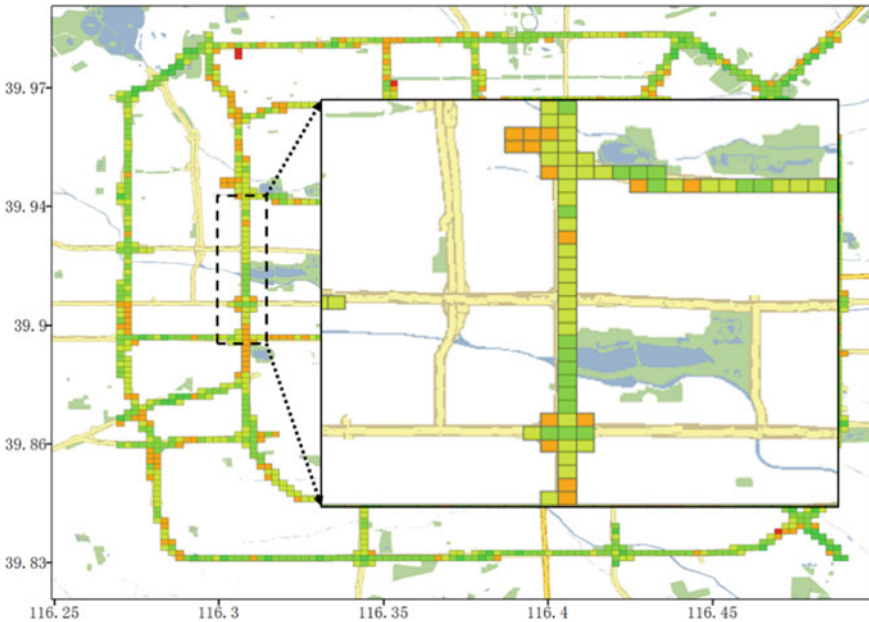


Fig. 2 The section of expressway selected

The experiment platform is server with 32 CPU cores (Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.1 GHz), 64G RAM, and GPU (NVIDIA GeForce RTX 2080). The experiment utilizes python 3.6.1 with Scikit-learn [20], Tensorflow [2], Keras [4] on Windows 10 for comparing the models.

3.2 Experimental Results

Figure 3 shows the training and testing curves of MSE with the number of epochs. From Fig. 3a–c, it can be seen that the iterative process curve flattens as the number of training epochs increases. This suggests that the convergence of LSTM and the convergence rate of LSTM are rapid. In addition, from Fig. 3d, it can be seen that the MSE of the 30-min time step is greater than others, and this can be attributed to the fact that the longer the time step, the more losses.

Figure 4 shows the prediction performance in the northbound direction on July 13 (Wednesday), 22 (Friday), and 24 (Sunday). It is obviously shown that the result of predicting the speed of the 10-min time step is the best, followed by 20 and 30 min. Meanwhile, there was no large difference for the speed trend between Wednesday’s

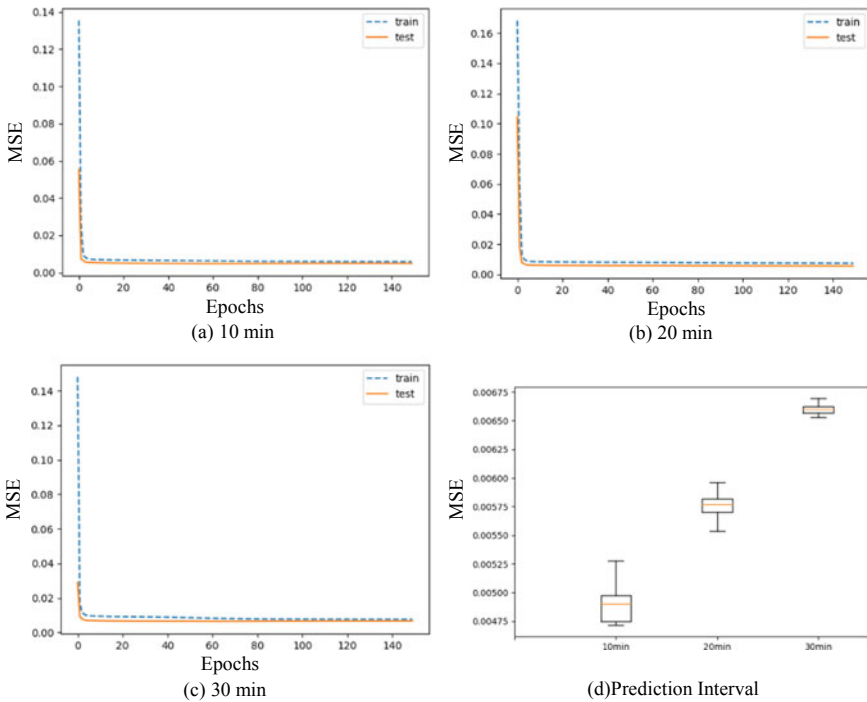


Fig. 3 Learning processes of LSTM

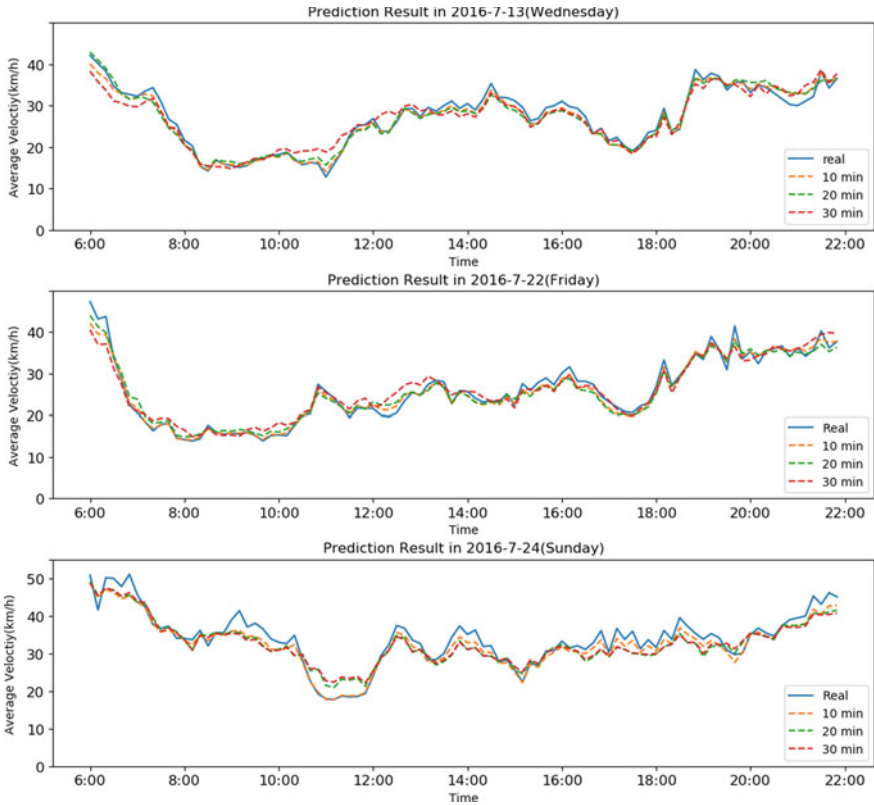


Fig. 4 Result of prediction of three days in test dataset

and Friday’s, whose peak hours were basically consistent. However, the morning peaks on the weekend came later than the weekday, and the evening peak hours on the weekend were also not obvious by comparing it with weekdays. Above all, when speed fluctuates greatly, the LSTM is capable of capturing the trend and fluctuation of the traffic speed in both the peak and off-peak hours.

In summary, Table 3 lists the six-day dataset (i.e. 864 sample data) of test dataset (i.e. 1728 sample data, accounted for 30% of the total dataset). We measure the accuracy as the Root Mean Square Error (RMSE) and R-Squared (R^2). It can be seen that the error eventually increases with the increase in estimation interval.

Figure 5 shows the prediction results of the corresponding average statistical tests (RMSE) of the proposed model. It can be found that RMSE fluctuates significantly during the 0:00–6:00 period, which is bigger than that in the daytime. The reason is that during the 0:00–6:00 period, the speed fluctuates greatly due to data miss. Although we use the exponential smoothing method to supplement the missing data, it still has a slight impact on the prediction results. Furthermore, due to the more stable

Table 3 Prediction accuracy in different time intervals

Date	10 min		20 min		30 min	
	RMSE	R ²	RMSE	R ²	RMSE	R ²
2016-7-13	0.933	0.991	2.326	0.946	3.321	0.887
2016-7-15	1.130	0.983	2.817	0.896	3.821	0.808
2016-7-17	1.096	0.983	3.526	0.809	3.871	0.770
2016-7-20	2.337	0.965	3.579	0.814	3.880	0.801
2016-7-22	0.933	0.995	2.751	0.934	3.741	0.880
2016-7-24	1.116	0.992	2.619	0.926	3.453	0.882

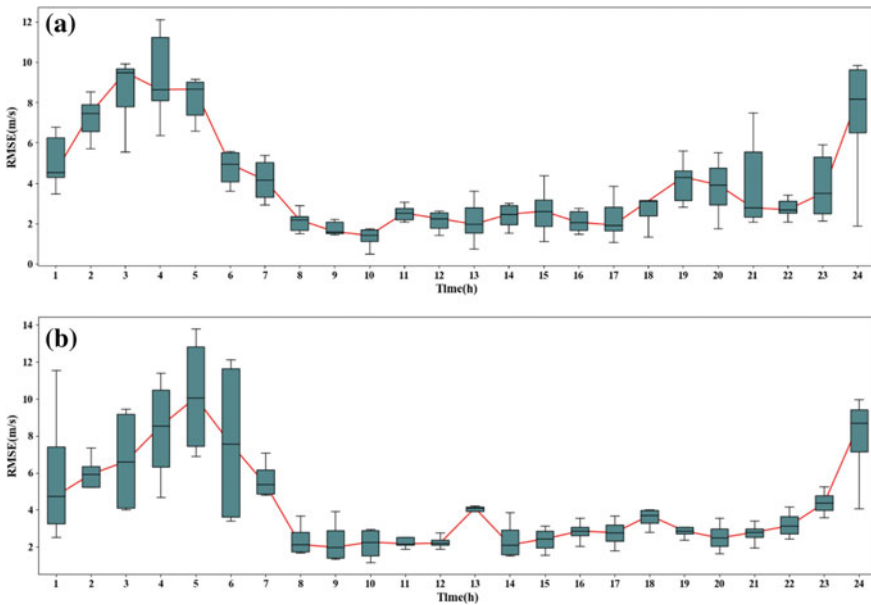


Fig. 5 RMSE of LSTM on FCD set during a day **a** on weekdays, **b** on weekends

speed fluctuation in the weekend evening peak period (19:00–21:00), the prediction performance is better than that on weekdays.

3.3 Comparison to Other Machine Learning Algorithms

We compare our model LSTM with other classical baseline statistical models that include K-Nearest Neighbor (KNN), Support Vector Regression (SVR) [25, 31] and Recurrent Neural Network (RNN).

Table 4 Predicted performance of LSTM and others

	10 min			20 min			30 min		
	MAE	RMSE	R ²	MAE	RMSE	R ²	MAE	RMSE	R ²
KNN (k = 16)	3.840	5.537	0.870	4.455	6.153	0.838	4.839	6.491	0.819
SVR	3.422	5.160	0.884	4.097	5.793	0.854	4.555	6.207	0.834
RNN	1.170	1.496	0.977	2.390	3.055	0.884	2.859	3.615	0.840
LSTM	1.044	1.424	0.979	2.239	2.861	0.901	2.771	3.544	0.846

LSTM model has higher prediction accuracy

Table 4 showed the MAE, RMSE and R² values for different time intervals. According to our results, LSTM had better accuracy in most of the cases compared to other models over 10, 20 and 30 min estimation intervals.

In order to check model’s accuracy in different time distribution, we measure the prediction accuracy during morning peak hours (7:00–10:00), evening peak hours (17:00–20:00) and off peak hours (14:00–16:00) on weekdays (see Fig. 6). It can be intuitively seen that the worst model is KNN. Although the median of RMSE of

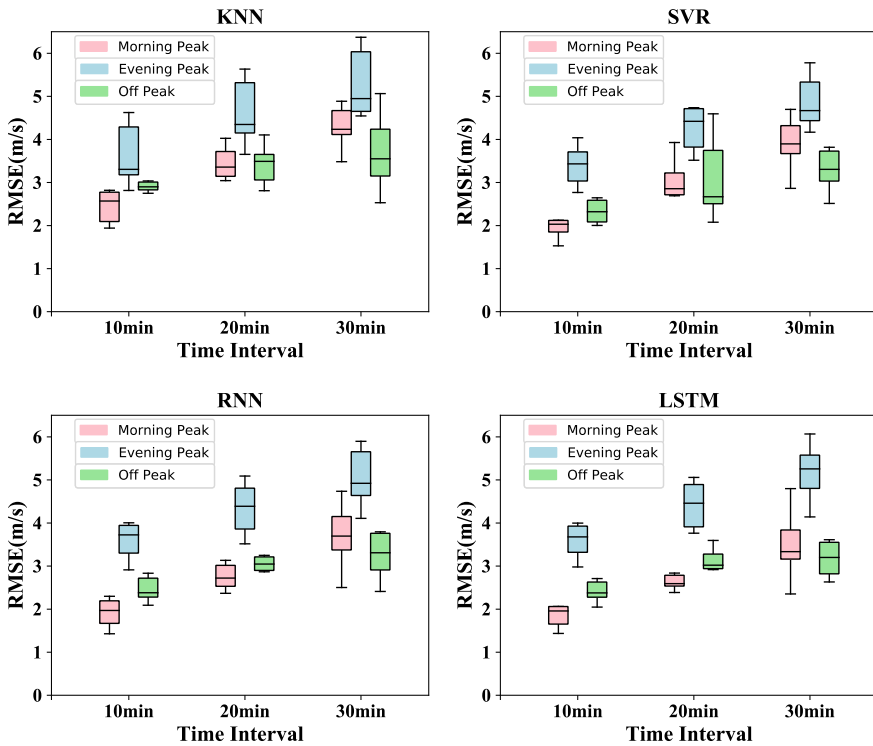


Fig. 6 Predicted performance of LSTM and others at peak hours and off peak hours

SVR is lower than that of KNN, the RMSE fluctuation of SVR is bigger. It is also noticed that the predicted performance of RNN is better than KNN and SVR, but compared with LSTM model, it is still slightly inferior.

4 Conclusion

This paper presents a novel long short-term memory neural network to predict travel speed using floating car data. To validate the effectiveness of the proposed LSTM, 40 days' traffic speed data with the updating frequency of 1 min from Beijing's expressway were collected. In order to enhance calculative efficiency, we use grid model to match trajectory data to road segments. Based on the grid model, a set of key traffic parameters (e.g. upstream average speed, downstream average speed, traffic volume, historical median speed, historical average speed, average speed) are derived as input data. The predicted speed datasets over 10, 20 and 30 min estimation intervals are output dataset. The seventy percent of the dataset was utilized for training, and the remaining was to test the algorithm performance. We demonstrated the usefulness and superiority of the proposed model by comparing it to other machine learning methods, the experimental results show that the MSE value obtained by LSTM is lower than that of the existing methods (KNN, SVR and RNN), which indicates that our method has a better forecasting performance.

Future work can be conducted by adding multiple layers into LSTM such as Stack-LSTM (i.e. Multilayer-LSTM), GRU (i.e. Multilayer-GRU), and Bidirectional LSTM (i.e. Multilayer-Bidirectional-LSTM). In addition, the prediction performances with different data aggregation levels should be investigated.

Acknowledgements This work is financially supported by Fundamental Research Funds for Beijing Jiaotong University (2019JBZ003). Many thanks are extended to the Institute of Beijing Taxi 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. Abbas Z et al (2018) Short-term traffic prediction using long short-term memory neural networks. In: 7th IEEE international congress on big data, big data congress 2018, 2 July–7 July 2018. Institute of Electrical and Electronics Engineers Inc.
2. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mané D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X (2015) TensorFlow: large-scale machine learning on heterogeneous systems. Software available from <http://tensorflow.org/>

3. Cai L et al (2020) A sample-rebalanced outlier-rejected k-nearest neighbour regression model for short-term traffic flow forecasting. *IEEE Access* 1–11
4. Chollet F et al (2015) Keras. <https://github.com/fchollet/keras>
5. Chen Y, Liu W, Zhang (2019) Evaluating and diagnosing road intersection operation performance using floating car data. *Sensors* 19(10):2256
6. Chen T, Lu S (2016) Accurate and efficient traffic sign detection using discriminative AdaBoost and support vector regression. *IEEE Trans Veh Technol* 65(6):4006–4015
7. Cui Z, Ke R, Wang Y (2018) Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction. 1801.02143
8. Deng B et al (2015) Estimating traffic delays and network speeds from low frequency GPS taxis traces for urban transport modelling. *Eur J Transp Infrastruct Res* 15(4)
9. Ehlers SFG (2019) Traffic queue length and pressure estimation for road networks with geometric deep learning algorithms. *Machine Learning*
10. Fabrizi V, Ragona R (2014) A pattern matching approach to speed forecasting of traffic networks. *Eur Transp Res Rev* 6(3):333–342
11. Fu R, Zhang Z, Li L (2016) Using LSTM and GRU neural network methods for traffic flow prediction. In: 2016 31st youth academic annual conference of Chinese association of automation (YAC). *IEEE*
12. Haworth J et al (2014) Local online kernel ridge regression for forecasting of urban travel times. *Transp Res Part C Emerg Technol* 46:151–178
13. He Z et al (2017) Mapping to cells: a simple method to extract traffic dynamics from probe vehicle data. *Comput-Aided Civ Infrastruct Eng* 32(3):252–267
14. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
15. Hua H, Jiang-Jie D (2012) Principle analysis and application of road traffic microwave sensor for highway. *Internet of things technologies*
16. Kim T-Y, Cho S-B (2018) Web traffic anomaly detection using C-LSTM neural networks. *Expert Syst Appl* 106:66–76
17. Leduc G (2008) Road traffic data: collection methods and applications. *Work Pap Energy Transp Clim Change* 1(55)
18. Ma X et al (2015) Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transp Res Part C Emerg Technol* 54:187–197
19. Nair V, Hinton GE (2010) Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*
20. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830
21. Tan YH, Chan CS (2017) Phrase-based image captioning with hierarchical LSTM model. *Comput Vis Pattern Recognit*
22. Turksma S (2000) Various uses of floating car data. In: *International conference on road transport information & control IET*
23. Vanajakshi L, Rilett LR (2004) A comparison of the performance of artificial neural networks and support vector machines for the prediction of traffic speed. In: *IEEE intelligent vehicles symposium, 2004*. *IEEE*
24. Van Lint JWC, Van Hinsbergen CPIJ (2012) Short-term traffic and travel time prediction models. *Artif Intell Appl Crit Transp Issues* 22(1):22–41
25. Wu C-H, Ho J-M, Lee D-T (2004) Travel-time prediction with support vector regression. *IEEE Trans Intell Transp Syst* 5(4):276–281
26. Wang X, Xu L, Chen K (2019) Data-driven short-term forecasting for urban road network traffic based on data processing and LSTM-RNN. *Arab J Sci Eng* 44(4):3043–3060
27. Yang B et al (2019) Traffic flow prediction using LSTM with feature enhancement. *Neurocomputing* 332:320–327
28. Yu B et al (2016) k-Nearest neighbor model for multiple-time-step prediction of short-term traffic condition. *J Transp Eng* 142(6):04016018

29. Yu R et al (2019) LSTM-EFG for wind power forecasting based on sequential correlation features. *Future Gener Comput Syst* 93:33–42
30. Yuan H, Tiandong XU, Lijun S (2008) Fuzzy logics based traffic state identification on urban expressway. *Highw Eng*
31. Zhang Y, Liu Y (2009) Traffic forecasting using least squares support vector machines. *Transportmetrica* 5(3):193–213
32. Zhao J, Mao X, Chen L (2019) Speech emotion recognition using deep 1D & 2D CNN LSTM networks. *Biomed Signal Process Control* 47:312–323

Design and Implementation of Automobile Test Monitoring Management System Based on PDA Terminal



Hong Jia, Hao Li, and Hai-ying Xia

Abstract In order to ensure the traceability of the test and detection work of road transport standard vehicles, and supervise and manage the test quality of automobile inspection institutions, this paper designs a monitoring and management system for automobile test and detection based on PDA terminal, and realizes the functions of PC terminal and PDA terminal by using C++ language and Java JDK1.8 language. Using PDA terminal, through real-time recording the operation data of the test sample vehicle or remote synchronous video recording, the original data of the test and detection of the standard vehicle are saved to the inspection quality supervision and management system, so as to ensure the traceability of the test and detection work of the standard vehicle type test organization, and on this basis, the test work quality of the standard vehicle type inspection organization is supervised and managed.

Keywords PDA terminal · Automobile test and inspection · Monitoring management system · System alarm

1 Introduction

In recent years, with the improvement of people's living standards, online shopping platforms have increased rapidly, logistics industry has developed rapidly, and road transport vehicle team has gradually grown up. With the rapid development of road transportation, the safety situation of road transportation in China is not optimistic. Although the safety level of road transportation has been improved year by year, in recent years, serious and serious traffic accidents with mass casualties still occur in our country, causing huge losses to the national economy and people's lives and property. Traffic accident cases show that in all kinds of accidents, operating vehicles

H. Jia (✉) · H. Xia

Research Institute of Highway Ministry of Transport, Ministry of Transport, Beijing, China
e-mail: h.jia@rioh.cn

H. Jia · H. Li · H. Xia

Key Laboratory of Operation Safety Technology on Transport Vehicles, Ministry of Transport, Beijing, China

are the main cause of traffic accidents on ZHONGTE Avenue. In addition to drivers, roads and climatic conditions, the technical performance of operating vehicles and mechanical faults in the use process are also the causes of traffic accidents on chongte Avenue [1].

At present, China has carried out a large number of specific technical services around the technical management of road transport vehicles, optimized the supervision measures for the technical management of operating vehicles, put forward the most basic safety technical requirements for the safety performance and structural configuration of road transport vehicles, and issued some standards and specifications for road transport vehicles test and detection. Therefore, it is necessary to establish a monitoring platform for road transport vehicle test, and strengthen the technical access and inspection supervision of road transport vehicles.

With the explosive development of science and technology, the performance of PDA (personal digital assistant) has been rapidly improved, and has been widely used in all walks of life [2–5]. Its biggest characteristic is that it has an open operating system, supports software and hardware upgrades, integrates information input, storage, management and transmission into one, and has common office, entertainment, mobile communication and other powerful functions. Therefore, PDA can be called a mobile office. This paper designs and implements an automobile test monitoring management system based on PDA. By recording the operation data or video of the test sample vehicle in real time, the system saves the original data of the standard vehicle test, ensures the traceability of the standard vehicle test and detection, and takes it as the basis for the test of the qualified vehicle test organization to supervise and manage the work quality.

2 System Architecture Design

The system adopts three-tier architecture: database access layer, server business logic layer and client side (Fig. 1).

The client design is as simple as possible, only dealing with the response and display of interface controls and the realization of audio and video. Check the data input to avoid data processing. In the design, the function is modularized, the interface of different functions is distinguished, and the coupling between different controls is reduced.

The user logic module is used to maintain the current login user information. Audio and video forwarding logic module is used to transfer media data and download video files between devices and clients. In the experiment video module, the audio and video data collected in the test process are combined into a general AVI file. GPS alarm module is used to analyze GPS data and follow the alarm principle. The data query module, according to the information uploaded by the vehicle terminal, establishes the database according to the vehicle VIN number, and can conduct comprehensive query according to the keywords reported by the test project, tester, test site and other detection institutions before the test. Message processing module provides various

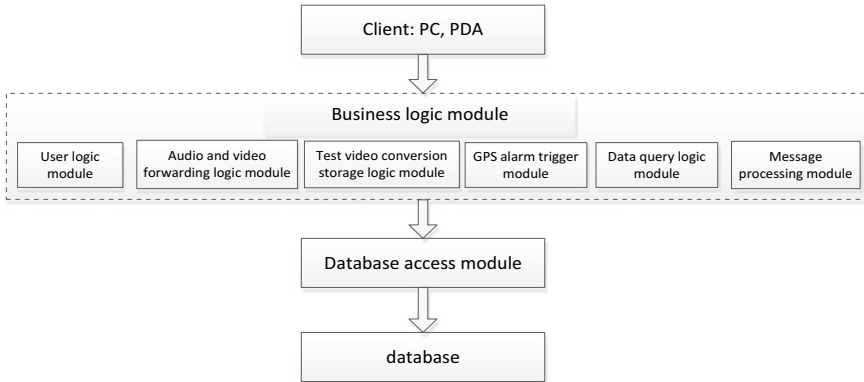


Fig. 1 System technical architecture

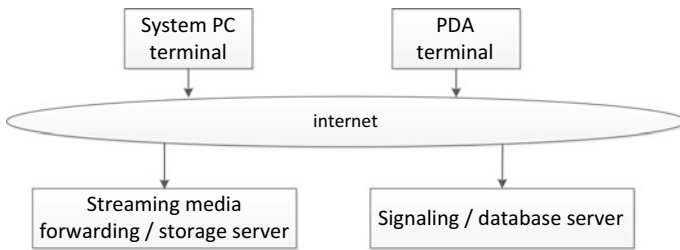


Fig. 2 Network service topology

processing methods, accepts data requests from UI, calls database access module for processing, and returns processing results to UI layer. Database access module, using powerful SQL Server 2005 database, encapsulates the operation of the database.

Add a layer of network service layer on the business logic layer to respond to the data request from the client on the network, verify its security, and return the calculated and processed data. The network service topology of the system is as in (Fig. 2).

Because of different positions in different institutions, the courses that all kinds of personnel need to learn are also different. Therefore, this paper classifies the target users, as shown in Table 1.

3 System Function Design

In order to better ensure the authenticity and effectiveness of automobile test and detection, this paper designs a set of automobile test and detection monitoring management system based on PDA terminal, which retains the video of key vehicle

Table 1 Test information

Serial number	Field name	Content
1	Name of declared enterprise	–
2	Name of testing organization	–
3	Test site	–
4	Test manager	–
5	Declared vehicle VIN	–
6	Vehicle model	–
7	Testing personnel	–
8	Vehicle monitoring terminal ID	–
9	Fuel consumption meter equipment number	–
10	Speedometer equipment number	–

test and detection items, and collects the track data through the PDA terminal of automobile test and detection. The database is established according to the vehicle VIN code, which can conduct comprehensive query and statistics according to the keywords reported by the testing institutions such as test items, testers and test sites before the test, so as to realize the storage and access function of uploaded video and track data. GIS geographic map can display the real-time satellite positioning information and track playback function of test vehicles, and automatically judge whether the road test meets the specifications according to the satellite positioning information.

3.1 Function Design of PC Terminal

The PC terminal of the system mainly includes user management, test video management, test information input, test process control, test data management, system alarm management, GPS real-time positioning and historical track playback.

- (1) **System User Management.**
System administrators can add and delete system users, and view the login status and log of users.
- (2) **Test video management.**
The system manager can monitor and view the video of single or multiple vehicles, regional setting, command issuing and other functions, and can monitor or view historical video data in real time. The system reviewer can check the video information and mark the video that does not meet the test specifications.
- (3) **Test information input.**
The testing organization shall input the test related information as follows.
- (4) **Test process control.**

Table 2 System alarm rules

Serial number	Alarm category	Set rules
1	Mileage alarm	If a certain vehicle from the beginning of the test to the end, if the middle mileage is less than 13 km, the alarm will be given
2	Time alarm	If the time between the beginning of the test and the end of a vehicle is less than 90 min, the alarm will be given
3	Terminal fault alarm	If the terminal uploads fault information, it will alarm
4	Test stop alarm	If the testing organization submits the test termination information, it will alarm
5	Regional alarm	If the vehicle drives out of the test area during the test, the alarm will be given

The corresponding test can be found, and the status of the test can be suspended, continued or voided.

(5) Test data management.

According to the information uploaded by the vehicle terminal, the database is established according to the vehicle VIN code, and the comprehensive query and statistics can be carried out according to the keywords reported by the testing institutions such as test items, testers and test sites before the test, so as to realize the storage and access functions of the uploaded data, such as adding, deleting, modifying and querying data.

(6) System alarm management.

The system can alarm according to mileage, time, terminal fault, test termination and area and other abnormal conditions. The rule settings are shown in Table 2.

The system can count the number of alarms according to time, testing organization, test vehicle type, vehicle VIN code. If there are more than two alarm messages for vehicles with the same VIN code, it will be highlighted.

(7) GPS real time positioning and historical track playback.

Display the current position of the vehicle while opening the real-time video. Display the historical track according to the test items, and click the icon on the map to display the longitude and latitude, speed, direction and time of the moment. The map has the basic operations, such as zoom in, zoom out, zoom in, zoom out, distance measurement, scale display, print and save the current screen image. The system has GIS geographic map, which can display real-time satellite positioning information and track playback function of test vehicles; according to the satellite positioning information, it can automatically judge whether the road test meets the specifications.

3.2 *Function Design of PDA Terminal*

The PDA terminal used for automobile test and detection is mainly used for shooting, retaining, positioning and playback of vehicle test related videos, including four functional modules: start test, end test, stop test and continue test.

- (1) Start the experiment.
After the user clicks the new test button, the user can create a new test interface module. After filling in the test vehicle type and other information according to the interface, the system judges the validity of the input information, calls the database access module to write the data, establishes the binding relationship between the test monitoring equipment and the VIN of the test vehicle, and updates the database by the database access module. Through the PDA terminal to shoot the relevant test project video, after uploading the video, the PC terminal can query the relevant content.
- (2) End the test.
The user can click the end test button to obtain the test information that has been started from the server, select the corresponding test, modify the status to end, and submit it to the server. The test control module calls the database access module to modify the data and release the binding relationship between the monitoring equipment and VIN, and the database access module updates the database.
- (3) Discontinue the test.
The user can click the stop test button to obtain the test information that has started from the server, create the test termination interface module, select the corresponding test modification status as abort and submit it to the server. The test control module calls the database access module to modify the data and release the binding relationship between the monitoring equipment and VIN.
- (4) Continue the experiment.
The user clicks the continue test button to obtain the information of the aborted test from the server. Create the interface module of continuous test, select the corresponding test modification monitoring equipment, and submit it to the server. The test control module calls the database access module to modify the data and establish the binding relationship between the monitoring equipment and VIN.

4 System Development and Implementation

The PC terminal of the system is developed with C++ language and VS2005 as compiler. At the same time, the system uses plug-in structure to facilitate function expansion, uses UDP high reliability transmission mechanism to ensure signaling, and has the advantages of accurate data and low delay. At the same time, the video uses H.264 hard decoding to reduce power consumption and make the picture more fluent.

PDA terminal software is developed based on Android architecture, the main programming language is java JDK1.8, which can be used in Android 7.0 or above system. The PDA terminal uses hardware to accelerate H.264 encoding and decoding, which is more power-saving and efficient. At the same time, it supports mainstream chip architectures such as arm, armv7a, arm64v8a, and is flexible to deploy. Low delay QOS transmission mechanism is adopted to ensure smooth and stable real-time video.

The system is deployed on two servers on Alibaba cloud. One server is used for streaming media forwarding or storage, and the other server is used for database storage. The two cloud servers are configured with 8-core CPU and 16 g memory. At the same time, Alibaba cloud security products such as web application firewall, situation awareness, and anknight are configured to ensure the data security of the system.

5 Concluding Remarks

Based on the analysis of user's functional requirements, this paper designs a monitoring and management system for automobile test and detection based on PDA terminal, and realizes the functions of PC terminal and PDA terminal by using C++ language and Java JDK1.8 language. Vehicle test and detection personnel can use PDA terminal to shoot test video, locate test vehicle and replay historical motion track. At the same time, each test vehicle has a unique VIN code. Therefore, through the test data, it is ensured that the vehicle type to which the test vehicle belongs can meet the requirements of various standards and specifications required by road transportation vehicles. The design and implementation of the system will help road transport vehicle technology access related work. With the promotion and application of the system, it will effectively improve the safety level of road transport industry and promote the transformation and upgrading of the industry.

Acknowledgements The project is supported by the scientific and technological innovation fund program of the Research Institute of Highway of the Ministry of Transport (2020-9067) and the scientific and technological innovation fund program of the Research Institute of Highway of the Ministry of Transport (2018-C0012).

References

1. Yuanpeng L, Xiaoping T, Shuquan X, Huiming N (2013) Thinking and countermeasures of technical management of operating vehicles. *Auto Maint Repair* 256(06):17-19
2. Chunmei L, Hongzhi L, Hongli Q (2020) Research on literature resources construction mode based on PDA. *Sci Technol Innov* 148(04):49-51
3. Cui L, Xiaohuan T (2019) Design of purchase, sale and stock management system for small and medium sized enterprises based on PDA. *Shipbuild Vocation Educ* 740(06):84-86

4. Tong Y, Qinghua Z, Ming C, Yanran Q, Jiajing G (2019) Design of automatic data acquisition platform for electric energy measurement based on PDA. *Autom Instrum* 241(11):119–122
5. Fengping Z, Yuanchun W (2019) Exploration and practice of new book borrowing service based on PDA—taking “new book borrowing room” of northeast petroleum university library as an example. *Inner Mongolia Sci Technol Econ* 435(17):151–152

A Three-layer Optimal Distribution Problem for Electric Vehicle Charging Stations



Di Chen, Xinyu Yu, Linghan Li, and Yuyang Zhou

Abstract Following the boost in technology in recent years, electric vehicles have flourished at an amazing rate and have become viable replacements for gasoline vehicles. Accordingly, an urgent need for optimal land use planning (LUP) for charging stations (CSs) has emerged. With the purpose of satisfying this demand, this study aims to solve the charging station distribution problem by determining the suitable number of chargers in a station, optimizing the distribution of CSs, and generating growth plans in lines with countries' regulations. First, queuing theory was used to determine the optimal number of chargers in a station; the effectiveness was measured by adopting China as an example. Second, in order to optimize the area distribution of CSs, graph theory was used to determine the shortest path using Floyd algorithm (FA); the aptness of the distribution of CSs was examined in parts of South Korea. Finally, analytic hierarchy process (AHP) was used to test the growth plan of the CS network in the U. S., South Korea, China, and Cote d'Ivoire, considering population density distributions, wealth distributions, and geographical conditions.

D. Chen · Y. Zhou (✉)
Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing
100124, China
e-mail: zyy@bjut.edu.cn

D. Chen
e-mail: chendi2366@163.com

D. Chen
Department of Civil and Environmental Engineering, University of California, Davis, CA 95616,
USA

X. Yu · L. Li
College of Automation, Beijing University of Technology, Beijing 100124, China
e-mail: xyy1511@163.com

L. Li
e-mail: lilinghan77@126.com

X. Yu
Department of Electrical and Electronic Engineering, The University of Manchester M139PL,
Great Manchester, UK

Keywords Electric vehicle · Land use planning · Shortest path problem · Charging station

1 Introduction

The potential of electric vehicles (EVs) as environmentally friendly alternatives for gasoline and diesel cars is expected to grow in the near future. Currently, the land use planning (LUP) of EV charging stations (CSs) are under development. However, means for determining the suitable number of chargers in a CS are limited, and current methods for identifying the area distribution of CSs are problematic. Without effective measures, the further advancement of EVs will be curtailed, and the shift to the use of electric cars will be delayed. Thus, a complete LUP is urgently needed. In this regard, diverse models are established to correlate critical factors to support the transition to EVs.

In recent years, several studies have been conducted in various aspects of EVs. Arias et al. [1] proposed a mixed-integer linear programming model to handle the robustness in multi-stage conjunct expansion planning of electrical distribution systems as well as the distribution of EV CSs. Chaudhari [2] considered the electricity price band allocation and used the hybrid optimization algorithm cost degradation model. Dong et al. [3] considered the influence of public charging facility arrangement involving traveled electric miles and implementing genetic algorithm to search the optimal (or suboptimal) positions of public charging sites. The distribution of EV parking lots was considered. Shafie-Khah et al. [4] proposed that the energy problem should be regarded as the main factor that influences the distribution of EV parking lots. A two-level model based on renewable energy was constructed and used for the evaluation of strategies for wind and solar photovoltaic implementation. He et al. [5] focused solely on the driving range of EVs to locate optimal sites of CSs with their proposed bi-level programming model. The upper level is aimed at enabling the site of CSs to reach optimization, whereas the lower level formulates a constraint so that a user equilibrium pattern of the traffic assignment problem can be solved.

Four approaches for CS placement were proposed by Lam et al.: iterative mixed-integer linear program, greedy approach, effective mixed-integer linear program, and chemical reaction optimization [6]. A multistage CS placement scheme based on customer behavior was presented by Luo et al. The nested logit model was used for the prediction of the charging demand of CSs. Bayesian game was applied for the optimal station placement at every planning stage [7]. Sadeghi-Barzani et al. implemented the genetic algorithm and the cost of station development was correlated with EV energy loss and electric grid, sites of electric substations, and roads in urban areas in CS placement was considered; they evaluated CS placement and sizing to determine the profitability of charging price [8]. Similarly, Simorgh et al. considered the cost of investment, connection, loss, and demand response by constructing an incentive-based demand response program and conducting handling operations with the particle swarm optimization algorithm [9]. A two-step screening approach,

which is concerned with the impacts on the environment and service radius of CS, was identified by Liu et al. [10]. Then, they used a modified primal–dual interior point algorithm to solve the optimal sizing problem, taking into account the overall cost. According to the survey conducted in about 100 communities in Beijing, China and the sales data of EVs, Wu and Yang figured out that insufficient and unreasonable CSs are the barriers for the development of EVs [11]. Guo and Zhao further introduced sustainability on the environment, economy, and social criteria through a fuzzy technique to avoid subjective judgment by decision makers [12].

Noori et al. presented an exploratory modeling method for studying uncertainties over the price of the life cycle, the cost of environmental deterioration, and water footprint of EVs in different grid regions in the U.S. [13]. Onat et al. took the environmental and the economic factors into account, and provided new framework for multi-criteria decision-making [14]. Dominguez-Jimenez et al. found that the seasons have significant impacts on charging patterns on the basis of the data in Boulder, the U.S. There was an inverse correlation between the charging time duration and temperature [15].

Other techniques and issues were examined for the CSs system. Cao et al. [16] presented a system containing predicted charging reservations and parking duration of EVs at CSs and designed a periodic updating mechanism for reservation. Deilami et al. [17] proposed a real-time smart load management control scheme that minimizes the overall price of grid energy loss. Khan et al. [18] improved the voltage quality and the harmonic distortion losses. They designed an integrated EV charging system and battery storage system to stabilize the voltage in CSs in the uncertainty of the wind and solar power. Tian et al. [19] focused on safety, where a multi-objective formulation was proposed to optimize the charging schedule. In terms of charging strategies, Moghaddam et al. [20] presented a smart charging method that provides various charging options for plug-in EVs and includes direct current fast charging and battery swapping. The queueing problem was also taken into account. Gusrialdi et al. [21] developed an approach for enabling queues to coordinate with CSs along highways. This approach can reduce the waiting time for charging. Corresponding distribution algorithm and distributed decision making policy were made and implemented in a certain highways in the U. S.

According to the preceding studies, diversified factors that may influence the distribution and placement of CSs have been considered. However, despite the inclusion of elements such as the cost of CS development and electricity, the loss of energy, the deterioration of environment, the security of chargers, the driving range of EVs, and the behavior of customers, diversity of different countries can be added as an essential factor. The emerging technique should also be combined with long term factors of LUP. However, these methods may not be readily available and suitable to different countries.

In this paper, the number, distribution and development of CSs are discussed. The paper compares the differences between countries, including geography, population density, and wealth distribution. The results can be expanded to large scale following three steps. First, the analysis begins on small scale. We take advantage of the Queueing Theory (QT) to construct a model to decide the optimal number of

chargers in each CSs. Second, the focus shifts to the overall situation, where Floyd algorithm (FA) is applied to decide the most suitable distribution of CSs. Third, with purpose of extending the applicability of the methods worldwide, we take notice of the diversity between different countries by giving different indexes on the three basic elements, namely, population density distributions, wealth distributions, and geographical conditions. This is accomplished with the use of analytic hierarchy process (AHP). The data are collected from four countries: U. S., South Korea, China, and Cote d'Ivoire, each of which has unique characteristics in population density distributions, wealth distributions, and geographical conditions needed to confirm our model and check our results.

The paper is organized as follows. In Sect. 2, three models are constructed, which begin with reliable assumptions to simplify the problem. Three subsections are included: first, QT is used; second, FA is applied; and third, the AHP model is established. In Sect. 3, the models are applied to four distinct countries. In Sect. 4, the results and effectiveness of the models are analyzed and the further improvements are recommended.

2 Problem Formulation and Models Proposal

2.1 Assumptions

To simplify the charging issue, several reasonable assumptions are made before starting the analysis.

1. All of the charging is to get electricity. The electricity is used to support EV's journey.
2. The EV arrival process in CS obeys the Poisson distribution.
3. There is no stay on the halfway to EV's destination caused by sudden events.

With these assumptions, specific models and solutions can be implemented.

2.2 Number of Charging Stations

In this section, we focus on the capacity of the CSs. Insufficient quantity of chargers in a station lead to queuing or significant idle costs, which prompt us to determine the balance between supply and demand in a given station. Therefore, the essential problem of the charging system is the service ability in the facility that means the number of chargers in the station. Based on QT, the optimal number of chargers is relevant to the arrival rate of EVs.

P_0 is defined as the probability that no EV arrives at the moment t , λ as the arrival rate, which is the number of EVs that arrive at the CS per unit time, μ as the average

service time of charging, C as the number of chargers, and ρ as the service intensity, that is, the arrival rate divided by the service time. The number of EVs waiting in line is defined as L_q , which can be calculated in Eqs. (1) and (2).

$$L_q = \frac{P_0 \cdot \left(\frac{\lambda}{\mu}\right)^C \cdot \rho}{C!(1 - \rho)^2} \tag{1}$$

$$\rho = \frac{\lambda}{C \cdot \mu}. \tag{2}$$

Given an arbitrary moment t , with the number of customers arriving at this moment n , then the probability of n EVs in the charging system can be described in Eq. (3).

$$P_n = \begin{cases} \left(\frac{\lambda}{\mu}\right)^n \frac{P_0}{n!} & (n = 0, 1, \dots, C) \\ \frac{\left(\frac{\lambda}{\mu}\right)^C}{C!C^{n-C}} P_0 & (n \geq C) \end{cases}. \tag{3}$$

For a certain system, the parameters λ , μ , C and P_0 are constants which can be taken from surveys. Then, a single objective optimization function can be established.

g is defined as the minimum number of chargers in a station because the cost of the building stations needs to be considered. Therefore, we have inequality $C \geq g$. With these limits, the optimal the number of chargers can be calculated and revised. Then, the need to optimize the number of EV chargers in CSs can be objectively met.

According to the definitions and equations mentioned earlier, the problem can be regarded as a single objective optimization model. h is defined as multiples that practical numbers exceed or go below the capacity of the station, which represents the number of EVs waiting in line goes beyond or under the capacity of the station.

$$\begin{aligned} & \min C \\ & \text{s.t.} \begin{cases} L_q = \frac{P_0 \times (\lambda/\mu)^c \times \rho}{C!(1-\rho)^2} \\ P_n(t) = \begin{cases} \left(\frac{\lambda/\mu}{n!} \times P_0\right) & (n = 0, 1, \dots, C) \\ \frac{(\lambda/\mu)^c}{C!C^{n-c}} \times P_0 & (n \geq C) \end{cases} \\ \rho = \frac{\lambda}{C \times \mu} \\ L_q \geq hC \\ C \geq g \\ C \in Z \end{cases} \end{cases}. \tag{4}$$

2.3 Allocation of Charging Stations

The allocation of CSs is one of the vital elements that influence the development of EVs. With a suitable allocation, EVs will develop at a quicker and smoother rate because of the convenience of charging. Thus, the optimization of the allocation of CSs is essential.

To optimize the allocation of CSs, FA is used as the major tool. The FA requires some nodes that are already known to calculate the shortest distance between any two nodes in the network. We choose the required nodes from scenic spots (SSs) and motorway service areas (MSAs) in this network. At SSs, people tend to stay for a long time, which enables them to charge their EVs. At same time, during long journeys in highways, EVs can only charge at MSAs. These two kinds of areas are suitable for building CSs because of the location and essential demand. Therefore, they are chosen as initial nodes to find the optimal location for the CSs. The shortest distance between each origin–destination pair of nodes is calculated and compared with the travel distance that a fully charged EV runs. If the distance in the FA goes beyond the remaining range of the fully-charged EVs, then we add nodes in between; if the calculated distance is too short that EVs can still go to the next CS before the electricity runs out, this CS is regarded as unnecessary and cancel it. Eventually, the remaining initial nodes and the added ones are the optimal positions where CSs should located. In this way, the optimal allocation of CSs can be obtained.

Several assumptions are considered in the model.

1. EVs choose the shortest route from the origin to the destination.
2. EVs only go to the CSs that are on their shortest route for charging.
3. The EV driver will not charge until there is no more than 50% electricity left.
4. After getting charged in the CSs, EVs will continue on to their destination.
5. The electricity consumption of EVs is relevant to the route distance.
6. The construction cost of CS in the road network is homogeneous, that means, the budget constraint is not related to the location of the CS, which is different from the wireless charging lane (WCL) system where the cost is due to the location and length of WCL.

The core of the optimal location model is to select the optimal position for the maximization of the benefit under budget constraint. For the charging station allocation problem, the goal is to satisfy the charging demand of EVs with the minimum consumption of the electricity on their way to the CS. Hence, the objective of the optimization model is to take the shortest path in the road network.

$G = (V, E)$ is used to define the network, with the node set V and the edge set E . We define two matrix first, namely, matrix S and P . Elements s_{ij} in matrix S represent the distance between node i and node j , where i and j is the serial number of nodes. Elements p_{ij} in matrix P represent the serial number of nodes that are on the way from i to j . We begin the FA with the following steps.

Step 1: s_{ij} is defined as the weight between i and j , which is the distance between i and j in this paper. If i and j are not adjacent, $s_{ij} = \infty$.

Step 2: The shortest distance between any two nodes is calculated to update matrix S and P . The matrix needs to be update N times, where N equals to the number of nodes in the graph. In the k th update, if $s_{ij} > s_{i(k-1)} + s_{(k-1)j}$, we replace s_{ij} to $s_{i(k-1)} + s_{(k-1)j}$. At the same time, $p_{ij} = p_{i(k-1)}$.

Step 3: After updating the matrix S and P for N times, the shortest path matrix (SPM) can be obtained, whose elements are the minimum distances between every two nodes.

After the three steps, the nodes necessary for building CSs are selected from the initial nodes. The goal is to build the minimum number of CSs and guarantee that all EVs in a chosen area get recharged before their electricity run out.

The choosing process is carried out with the following steps.

First, the mileage that 50% electricity can cover is defined as G (in kilometers).

Second, the elements that are greater or equal to G are selected from the SPM, and name them as a_{mk} , where $m, k = 1, 2, \dots, n$, n is the order of the SPM.

Third, all nodes on the shortest path are traversed from m . Then, accumulate the distance of the passing node on the shortest path and find the node closest to $\frac{a_{mk}}{2}$. This node is needed to build the CS. If there are two adjacent nodes that have the same distance to $\frac{a_{mk}}{2}$, then the one that is closer to the starting node is selected as the CS.

Last, the selected nodes are considered as the union. In this way, the optimal allocation of CSs can be obtained, which not only has the minimum number of CSs but also meets the charging demand.

2.4 Growth Plan of Charging Stations

The electric car industry is undergoing development. In this part, we discuss how the CS network will enlarge in the growth of the electric car production. The number of chargers and distribution of the CSs discussed in the previous sections are general approaches. Each country has its own distinct characteristics that can be attributed to the economy, population, and geography [22, 23], these three elements of diversity are considered to plan the growth. In this case, AHP is adopted to facilitate the decision of the different countries for their unique CS planning schemes.

The analytic hierarchy is built according to the next four steps of AHP:

1. The trigger to the growth of the EV network is determined. The factors that affect the decisions and solutions are figured out.
2. The relationship between the factors is analyzed and a hierarchy structure is constructed.
3. In accordance with one of the upper standards, the levels of importance between factors are compared in the same level and the comparison matrix is built.
4. The relative weigh is calculated on the basis of the judgement matrix. The consistency is checked to demonstrate the feasibility of the classification system.

Without loss of generality, several assumptions are made as follows.

1. Each factor in the decision level has no relationship and impact to one other.
2. Countries make their decisions on the EV development project based on their own conditions.

Combining the target and related elements together, each layer, the contents, and the hierarchy structure are decided. In the target layer, the major goal is to choose the best solution in line with the national conditions. In the decision level, three elements are set in the deciding process, namely, geographical conditions, population density distributions, and wealth distributions. Three scenarios are considered in the scheme layer to build the urban chargers first, to build the rural chargers first, or to build the urban and rural chargers simultaneously.

To determine CSs developing scheme of a country, we involve three typical characteristics, the gross domestic product per capita (GDPPC), the population density (PD), as well as the national land area (NLA). The GDPPC, which measures the average economic status of a country, acts as a measurement of wealth distribution. The PD indicates the number of people in a unit land area, representing the population density distribution. The geographic situation of a country, which is relevant to the transportation network design and construction are indicated by NLA. The structure of the decision framework is shown in Fig. 1.

The weight of each pair of factors was calculated in Eq. (5).

$$w_i = \frac{1}{n} \sum_{j=1}^n \frac{x_{ij}}{\sum_{k=1}^n x_{kj}} \quad i = 1, 2, \dots, n, \tag{5}$$

where a_{ij} is the comparison score between factors i and j . n is the order of the matrix. Accordingly, we have matrix W consist of w_i . Moreover, x_{ij} is the reciprocal of a_{ji} , which is represented by Eq. (6).

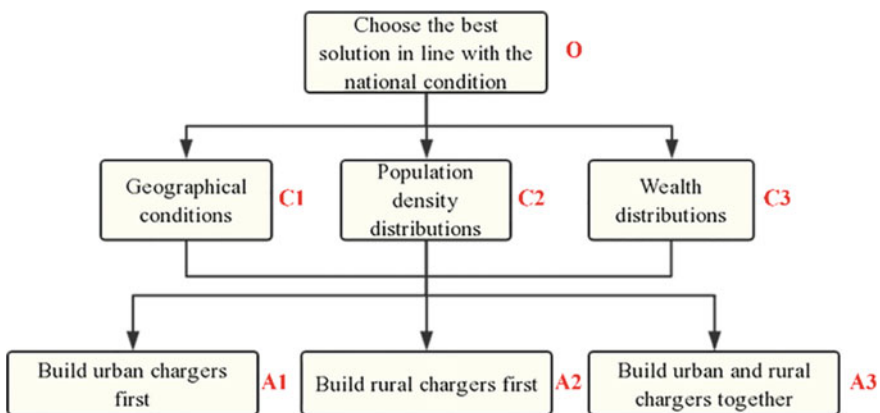


Fig. 1 Analytic hierarchy process of electric vehicle distributions

$$x_{ij} = \frac{1}{x_{ji}}. \tag{6}$$

In order to indicate the importance of the elements between i and j ranks from the vital to the inessential, there are nine values of x_{ij} , namely, $\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, \frac{1}{1}, \frac{3}{1}, \frac{5}{1}, \frac{7}{1}, \frac{9}{1}$. We compare each two factors and decide their level of importance. Then, matrix X can be obtained with x_{ij} as its values.

The matrix X is normalized as \bar{X} . The elements in \bar{X} is defined as \bar{x}_{ij} . The normalization follows Eq. (7).

$$\bar{x}_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}. \tag{7}$$

w_i is normalized as \bar{w}_i following Eq. (8).

$$\bar{w}_i = \frac{\sum_{j=1}^n \bar{x}_{ij}}{\sum_{i=1}^n \sum_{j=1}^n \bar{x}_{ij}}. \tag{8}$$

The consistency check is carried out. The eigenvalues of matrix X is calculated from Eq. (9).

$$\lambda_{\max} = \max \left\{ \frac{(X \cdot W)_i}{W_i} \right\}. \tag{9}$$

Then, the consistency index C.I. is calculated in Eq. (10).

$$C.I. = \frac{\lambda_{\max} - n}{n - 1}. \tag{10}$$

R.I. is the Random Index which is constant. The consistency ratio C.R. is calculated as Eq. (11).

$$C.R. = \frac{C.I.}{R.I.}. \tag{11}$$

The total weight from each layer function to the upper layer is calculated.

To make our calculation clearer, O represents the target layer. $C1$, $C2$, and $C3$ equals to geographical conditions, population density distributions, and wealth distributions respectively, which are in the decision level. $A1$, $A2$, and $A3$ are corresponding to building urban chargers first, building rural chargers first, as well as building urban and rural chargers simultaneously. Under different targets in different layers, we have divergent values of x_{ij} . This means that we get different \bar{w}_i following the previous calculation. \bar{w}_i is used as the measurement of the choice of CS construction plan. If O is the target, \bar{w}_i of O act as the index of comparison that we define as

\overline{w}_{i0} . Consequently, if C1, C2, and C3 are the target, we define \overline{w}_i as \overline{w}_{ik} , $k = 1, 2, 3$, respectively. Then, the overall priority index z_i was calculated according to Eq. (12). We choose z_i with the maximum value. Then, target that this value corresponds to is the most suitable construction plan.

$$z_i = \sum_{k=1}^3 \overline{w}_{i0} \times \overline{w}_{ik}. \quad (12)$$

3 Results and Discussion

3.1 Number of Chargers in China

To ascertain the number of chargers in each CS in China, we set $h = 2$, that is, the number of EVs waiting in line does not exceed 2 times of the station's capacity. Consequently, we have inequality $L_q \geq 2C$. Base on practical analysis in China, there should be at least 6 chargers in a station, thus $g = 6$. Under this circumstance, we get $C \geq 14$.

According to related studies [24], the average number of EVs arriving in 60 min was approximately 24.46 and with an average charging time of 30 min, so $\mu = 30$, $\lambda = 0.41$. The optimal number of chargers in each CSs is 19.

3.2 Distribution Performance in Parts of South Korea

To figure out the suitable distribution of CSs, a region of interesting South Korea is taken as case study. The region includes Seoul, the capital of South Korea, Incheon, which has an international airport, and Gyonggi Province, which surrounds Seoul and Incheon [25].

Following the principles mentioned in the previous section, the initial candidate nodes for the CS are chosen based on the service zones along the highway and the SSs, which are represented by the yellow dots and the green dots in Fig. 2.

Based on the graph theory, we establish a network by extracting the nodes in Fig. 2 and measuring the actual travel distances between related nodes as the weight, which is shown in Fig. 3.

The shortest path between every two nodes is determined by the steps of FA discussed in the previous chapter. The results are shown in the following matrix:

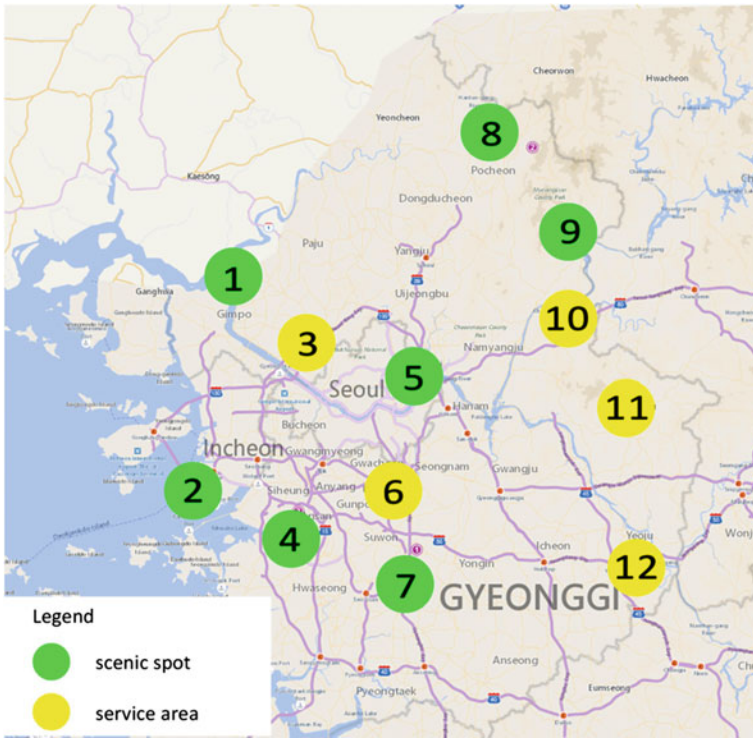
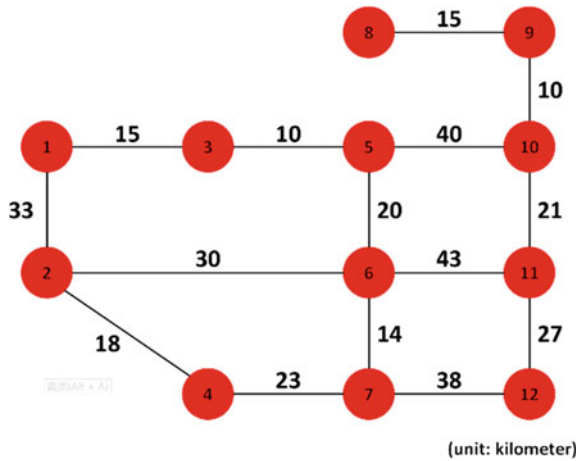


Fig. 2 Distribution of electric vehicle charging stations in three regions of South Korea

Fig. 3 Initial nodes and weights of electric vehicle charging stations in part of South Korea according to the Floyd algorithm



$$\begin{matrix}
 & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{7} & \mathbf{8} & \mathbf{9} & \mathbf{10} & \mathbf{11} & \mathbf{12} \\
 \mathbf{1} & 0 & 33 & 15 & 51 & 25 & 45 & 59 & 90 & 75 & 65 & 86 & 97 \\
 \mathbf{2} & & 0 & 48 & 18 & 50 & 30 & 41 & 115 & 100 & 90 & 73 & 79 \\
 \mathbf{3} & & & 0 & 66 & 10 & 30 & 44 & 75 & 60 & 50 & 73 & 82 \\
 \mathbf{4} & & & & 0 & 57 & 37 & 23 & 122 & 107 & 97 & 80 & 61 \\
 \mathbf{5} & & & & & 0 & 20 & 34 & 65 & 50 & 40 & 61 & 88 \\
 \mathbf{6} & & & & & & 0 & 14 & 89 & 74 & 64 & 43 & 70 \\
 \mathbf{7} & & & & & & & 0 & 99 & 84 & 86 & 65 & 38 \\
 \mathbf{8} & & & & & & & & 0 & 15 & 25 & 46 & 73 \\
 \mathbf{9} & & & & & & & & & 0 & 10 & 31 & 58 \\
 \mathbf{10} & & & & & & & & & & 0 & 21 & 48 \\
 \mathbf{11} & & & & & & & & & & & 0 & 27 \\
 \mathbf{12} & & & & & & & & & & & & 0
 \end{matrix} \tag{13}$$

According to the official data published by Renault Samsung Motors, we choose the SM3 ZE type of EV as our sample, where a fully charged EV can run 213 km [26]. To simplify our calculation, we assume that the distance that a fully-charged EV travel is 200 km. Accordingly, an EV with 50% electricity can run 100 km. Following the SPM calculated above, the choosing process is carried out. The result shows that there is only one node, No. 5, necessary to construct the CS.

3.3 Growth Plan of Countries

Different countries vary dramatically in their wealth distribution, population density distribution, and geography. The difference between countries may lead to distinct development schemes. Specifically, we choose four typical countries in our case study, including the U.S., South Korea, China, and Cote d’Ivoire. Since each of them has unique characteristics, the CSs’ construction strategy in each of them is worth analyzing. From the respective national websites, we get the data of GDPPC, PD, and NLA (Table 1), where the four countries are listed in the order of their GDPs.

The datasheet highlights the major characteristics of the four countries. The U.S. has the highest GDPPC among the countries. In South Korea, the major element

Table 1 Gross domestic product per capita, population density and national land area of four countries

Country	GDP (\$)	Population density (person/km ²)	National land area (km ²)
The U. S.	56,810	35.3	9,370,000
South Korea	27,539	525.7	100,284
China	8438	144.3	9,634,057
Cote d’Ivoire	1526	74.5	322,463

that influences the growth plan is the population density distribution. The reason is that South Korea has the highest PD and lowest NLA. China’s key factors include the wealth distribution and the population density distribution, due to its large NLA and high population density. Cote d’Ivoire is very special in terms of geographical condition because its NLA is relatively much smaller than that of the U. S. and China. The selection of the typical characteristics of counties is marked in Table 2.

A standard judgement matrix (SJM) is made for each country, which is given by Eq. (13). The basis of the judgment is the potential factors that influence the initial region of developing EV network, including wealth distributions, population density distributions, and geographical conditions. We compare these three factors with each other, deciding the level of importance and getting proper values from the nine values discussed in the previous chapter.

$$\begin{array}{cc}
 \begin{matrix} 1 & \frac{1}{3} & \frac{1}{7} \\ 3 & 1 & 5 \\ 7 & \frac{1}{5} & 1 \end{matrix} & \begin{matrix} 1 & 7 & 5 \\ \frac{1}{7} & 1 & 5 \\ \frac{1}{5} & \frac{1}{5} & 1 \end{matrix} \\
 \text{(a) the U.S.} & \text{(b) South Korea}
 \end{array} \tag{14}$$

$$\begin{array}{cc}
 \begin{matrix} 1 & 3 & 1 \\ \frac{1}{3} & 1 & \frac{1}{3} \\ 1 & 3 & 1 \end{matrix} & \begin{matrix} 1 & 3 & 5 \\ \frac{1}{3} & 1 & \frac{1}{3} \\ \frac{1}{5} & 3 & 1 \end{matrix} \\
 \text{(c) China} & \text{(d) Cote d'Ivoire}
 \end{array}$$

Note that in this problem, the order n of our matrix is 3. There are corresponding relationships between the order n and the value of R.I., which can be seen in Table 3. In this problem, the value of R.I. is 0.58.

R.I is the short for random index.

Table 2 Typical characteristics of each country

Country	Wealth distribution	Population density distribution	Geographical conditions
The U. S.	✓		
South Korea		✓	
China	✓		✓
Cote d’Ivoire			✓

Table 3 Values of random Index

Order	1,2	3	4	5	6	7	8
R.I	0.00	0.58	0.90	1.12	1.26	1.36	1.41

Table 4 Overall priority vector of four countries

	O	C1	C2	C3	Overall priority	Selection of scheme
		0.6234	0.1376	0.3099		
The U. S.	A1	0.1429	0.2000	0.6965	0.3475	
	A2	0.4286	0.2000	0.2368	0.2367	
	A3	0.4286	0.6000	1.0000	0.7050	✓
South Korea	A1	0.7235	0.6333	0.5558	0.6882	✓
	A2	0.0833	0.1062	0.3623	0.1131	
	A3	0.1932	0.2605	1.0000	0.2796	
China	A1	0.6333	0.4286	0.6965	0.6312	✓
	A2	0.1062	0.1429	0.2368	0.1674	
	A3	0.2605	0.4286	1.0000	0.6014	
Cote d'Ivoire	A1	0.6965	0.4286	0.6333	0.6446	✓
	A2	0.0719	0.1429	0.2895	0.1336	
	A3	0.2316	0.4286	1.0000	0.4423	

Under these prerequisites, the AHP is utilized to select the most suitable scheme for each country to build CSs. The results are shown in Table 4.

Using the data in Table 4, we obtain each country’s development plan for CSs, whether in urban or rural areas. To sum up, growth plan suggested for the four countries are shown in Table 5.

For the U. S., constructing city and rural CSs at the same time is preferable. By contrast, South Korea, China, and Cote d’Ivoire must divide their construction strategies into two phases, building CSs in cities first. The result also indicates that none of these countries should construct CSs in rural areas. The reason is that people tend to gather in urban areas, where they have higher EV ownerships and more urgent need for CSs. Under this circumstance, most countries around the globe should think about building their CSs at cities or simultaneously building CSs at cities and rural areas. Initial construction in rural zones is unfavorable, despite the differences among countries with respect to wealth distribution, population density distribution, and geography.

Table 5 Electric vehicle charging station developing strategies of studied countries

Country	Build urban chargers first	Build rural chargers first	Build urban and rural chargers together
The U. S.			✓
South Korea	✓		
China	✓		
Cote d’Ivoire	✓		

4 Conclusions and Future Works

The number of EVs has increased in many societies possibly because of the vast improvement in technology and increase in environmental awareness. There is an urgent need in optimal CSs' distribution plan. In this paper, we use three models to solve the EV charging problem. We start by calculating the optimal number of chargers in CSs, using the QT. Confirmation is carried out, and China is taken as an example. The result shows the most suitable number of chargers in each CS is 19 in China. According to our case study in parts of South Korea, CS distribution can be efficiently determined through FA. The initial location of chargers can be decided by MSAs and SSs and can then be adjusted by comparing the distance that fully-charged EVs can cover between two stations. Lastly, CSs' growth plan applicable to different countries are proposed with the use of AHP. We provide suggestions regarding countries' development plans in terms of constructing CSs in cities and rural areas. We recommend that distinct characteristics in countries' geographical conditions, population density distributions, and wealth distributions be considered during CS construction.

In future studies, the swapping of batteries in CSs is an expected problem in bus and taxi charging. The service rate parameter should be revised such that the number of chargers in each station can be calculated. Moreover, the weather conditions of a country, particularly during winter, may be another factor that influences the development of charging systems.

Acknowledgements The authors would like to thank the Consortium for Mathematics and Its Application of the United States for raising this problem. The authors appreciate the assistance from Prof. Zhengbing He and Mr. Minhe Zhao from Beijing University of Technology for the language edit.

Funding The work described in this paper was supported by the Natural Science Foundation of China (No. 71871010) and Beijing Municipal Education Commission, the General Program of Science and Technology (No. KM201810005017).

Appendix

Problem Raising by the Consortium For Mathematics and its Application of the United States

Problem D: Out of gas and driving on E (for electric, not empty).

For both environmental and economic reasons, there is global interest in reducing the use of fossil fuels, including gasoline for cars. Whether motivated by the environment or by the economics, consumers are starting to migrate to electric vehicles. Several countries are seeing early signs of the potential for rapid growth in the adoption of electric vehicles. In the US and other countries, the release of the more affordable all-electric Tesla Model 3 has resulted in record numbers of pre-orders

and long wait lists (<https://www.wired.com/story/tesla-model-3-delivery-timeline/>). To further accelerate the switch to electric vehicles, some countries, including China, have announced that they will ban gasoline and diesel cars in the coming years.

Eventually, when a ban goes into effect, there needs to be a sufficient number of vehicle charging stations in all the right places so that people can use their vehicles for their daily business, as well as make occasional long-distance trips. The migration from gasoline and diesel cars to electric vehicles, however, is not simple and can't happen overnight. In a fantasy world, we would wake up one day with every gas vehicle replaced by an electric one, and every gas station replaced with a charging station. In reality, there are limited resources, and it will take time for consumers to make the switch. In fact, the location and convenience of charging stations is critical as early adopters and eventually mainstream consumers volunteer to switch (<http://www.govtech.com/fs/Building-Out-Electric-Vehicle-Infrastructure-Where-Are-the-Best-Locations-for-Charging-Stations.html>).

As nations plan this transition, they need to consider the final network of charging stations (the number of stations, where they will be located, the number of chargers at the stations, and the differences in the needs of rural areas, suburban areas, and urban areas), as well as the growth and evolution of the network of charging stations over time. For example, what should the network look like when electric vehicles represent 10% of all cars, 30% of all cars, 50% of all cars, and 90% of all cars?

As nations seek to develop policies that promote the migration towards electric vehicles, they will need to design a plan that works best for their individual country. Before they can begin, they would like your team's help in determining the final architecture of the charging network to support the full adoption of all-electric vehicles. Additionally, they would like you to identify the key factors that will be important as they plan their timeline for an eventual ban or dramatic reduction of gasoline and diesel vehicles. To help your team manage the scope of this problem, we ask that you focus only on personal passenger vehicles (i.e. cars, vans, and light trucks used for passengers). At the end of your report, you may briefly comment on the relevance of your findings and conclusions on commercial vehicles to include heavy trucks and busses. Your tasks are the following:

Task 1: Explore the current and growing network of Tesla charging stations in the United States. Tesla currently offers two types of charging stations: (1) destination charging designed for charging for several hours at a time or even overnight (<http://www.tesla.com/destination>) and (2) supercharging designed for longer road trips to provide up to 170 miles of range in as little as 30 min of charging (<https://www.tesla.com/supercharger>). These stations are in addition to at-home charging used by many Tesla owners who have a personal garage or a driveway with power. Is Tesla on track to allow a complete switch to all-electric in the US? If everyone switched to all-electric personal passenger vehicles in the US, how many charging stations would be needed, and how should they be distributed between urban, suburban, and rural areas?

Task 2: Select one of the following nations (South Korea, Ireland, or Uruguay).

2a. Determine the optimal number, placement, and distribution of charging stations if your country could migrate all their personal passenger vehicles to all-electric vehicles instantaneously (no transition time required). What are the key factors that shaped the development of your plan?

2b. While these countries have already started installing chargers, you get to start with a clean slate. Present a proposal for evolving the charging network of your chosen country from zero chargers to a full electric-vehicle system. How do you propose the country invest in chargers? Should the country build all city-based chargers first, or all rural chargers, or a mix of both? Will you build the chargers first and hope people buy the cars, or will you build chargers in response to car purchases? What are the key factors that shaped your proposed charging station plan?

2c. Based on your growth plan, what is the timeline you propose for the full evolution to electric vehicles in your country? To get started, you may wish to consider how long it will take for there to be 10% electric vehicles, 30% electric vehicles, 50% electric vehicles, or 100% electric vehicles on your selected country's roads. What are the key factors that shaped your proposed growth plan timeline?

Task 3: Now consider countries with very different geographies, population density distributions, and wealth distributions, such as Australia, China, Indonesia, Saudi Arabia, and Singapore. Would your proposed plan for growing and evolving the network of chargers still apply to each of these countries? What are the key factors that trigger the selection of different approaches to growing the network? Discuss the feasibility of creating a classification system that would help a nation determine the general growth model they should follow in order for them to successfully migrate away from gasoline and diesel vehicles to all electric cars.

Task 4: The technological world continues to change and is impacting transportation options such as car-share and ride-share services, self-driving cars, rapid battery-swap stations for electric cars, and even flying cars and a Hyper-loop. Comment on how these technologies might impact your analyses of the increasing use of electric vehicles.

Task 5: Prepare a one-page hand-out written for the leaders of a wide range of countries who are attending an international energy summit. The hand-out should identify the key factors the leaders should consider as they return to their home country to develop a national plan to migrate personal transportation towards all-electric cars and set a gas vehicle-ban date.

Your submission should consist of:

One-page Summary Sheet,

One-page hand-out,

Your solution of no more than 20 pages, for a maximum of 22 pages with your summary and hand-out.

Note: Reference list and any appendices do not count toward the 22-page limit and should appear after your completed solution.

References

1. Arias NB, Tabares A, Franco JF, Lavorato M, Romero R (2018) Robust joint expansion planning of electrical distribution systems and EV charging stations. *IEEE Trans Sustain Energy* 9(2):884–894
2. Chaudhari K (2018) Hybrid optimization for economic deployment of ESS in PV-integrated EV charging stations. *IEEE Trans Ind Inf* 14(1):106–116
3. Dong J, Liu CZ, Lin ZH (2014) Charging infrastructure planning for promoting battery electric vehicles: an activity-based approach using multiday travel data. *Transp Res Part C Emerg Technol* 38:44–55
4. Shafie-Khah M, Siano P, Fitiwi DZ, Mahmoudi N, Catalão JPS (2018) An innovative two-level model for electric vehicle parking lots in distribution systems with renewable energy. *IEEE Trans Smart Grid* 9(2):1506–1520
5. He J, Yang H, Tang TQ, Huang HJ (2018) An optimal charging station location model with the consideration of electric vehicle's driving range. *Transp Res Part C Emerg Technol* 86:641–654
6. Lam AYS, Leung YW, Chu X (2014) Electric vehicle charging station placement: formulation, complexity, and solutions. *IEEE Trans Smart Grid* 5(6):2846–2856
7. Luo C et al (2015) A consumer behavior based approach to multi-stage EV charging station placement. Presented at VTC Spring. [Online]. Available: <https://ieeexplore.ieee.org/stamp.jsp?tp=&arnumber=7145593>
8. Sadeghi-Barzani P, Rajabi-Ghahnavieh A, Kazemi-Karegar H (2014) Optimal fast charging station placing and sizing. *Appl Energy* 125:289–299
9. Simorgh H, Doagou-Mojarrad H, Razmi H, Gharehpetian GB (2018) Cost-based optimal siting and sizing of electric vehicle charging stations considering demand response programmes. *IET Gener Transm Distrib* 12(8):1712–1720
10. Liu Z, Wen F, Ledwich G (2013) Optimal planning of electric-vehicle charging stations in distribution systems. *IEEE Trans Power Del* 28(1):102–110
11. Wu S, Yang Z (2020) Availability of public electric vehicle charging pile and development of electric vehicle: evidence from China. *Sustainability* 12(16):6369
12. Guo S, Zhao H (2015) Optimal site selection of electric vehicle charging station by using fuzzy TOPSIS based on sustainability perspective. *Appl Energy* 158:390–402
13. Noori M, Gardner S, Tatari O (2015) Electric vehicle cost, emissions, and water footprint in the United States: development of a regional optimization model. *Energy* 89:610–625
14. Onat NC, Noori M, Kucukvar M, Zhao Y, Tatari O, Chester M (2017) Exploring the suitability of electric vehicles in the United States. *Energy* 121:631–642
15. Dominguez-Jimenez JA, Campillo JE, Montoya OD, Delahoz E, Hernández JC (2020) Seasonality effect analysis and recognition of charging behaviors of electric vehicles: a data science approach. *Sustainability* 12(18):7769
16. Cao Y, Wang T, Kaiwartya O, Min G, Ahmad N, Abdulla AH (2018) An EV charging management system concerning drivers, trip duration and mobility uncertainty. *IEEE Trans Syst Man Cybern Syst* 48(4):596–607
17. Deilami S, Masoum AS, Moses PS, Masoum MAS (2011) Real-time coordination of plug-in electric vehicle charging in smart grids to minimize power losses and improve voltage profile. *IEEE Trans Smart Grid* 2(3):456–467
18. Khan A, Memon S, Sattar TP (2018) Analysing integrated renewable energy and smart-grid systems to improve voltage quality and harmonic distortion losses at electric-vehicle charging stations. *IEEE Access* 1–1
19. Tian A, Li WX, Li ZY, Sun Y (2018) Network security-aware charging of electric vehicles. *Int J Electr Power Energy Syst* 100:42–49
20. Moghaddam Z, Ahmad I, Habibi D, Phung QV (2018) Smart charging strategy for electric vehicle charging stations. *IEEE Trans Transp Electrif* 4(1):76–88
21. Gusrialdi A, Qu Z, Simaan MA (2017) Distributed scheduling and cooperative control for charging of electric vehicles at highway service stations. *IEEE Trans Intell Transp Syst* 18(10):2713–2727

22. Hu X, Wu C, Wang J, Qiu R (2018) Identification of spatial variation in road network and its driving patterns: economy and population. *Reg Sci Urban Econ* 71:37–45
23. Zhang C, Liu Y, Qiao H (2015) An empirical study on the spatial distribution of the population, economy and water resources in Northeast China. *Phys Chem Earth* 79–82:93–99
24. Lin L, Gao YJ (2014) Optimal allocation of charging stations based on queuing theory. *Electr Power Sci Eng* 30(4)
25. Bing map (Accessed on 19 September 2020). <https://cn.bing.com/destinationmaps?cid=7446&name=%E4%BA%AC%E7%95%BF%E9%81%93&lat=38.2223014831543&lon=127.011497497559&FORM=DESPCB>
26. Renault Samsung Motors (2018) (Accessed on 16 September 2020). <https://www.renaultsamsung.com/2017/main/main.jsp>

Analysis of Mixed Vehicle Traffic Flow at Signalized Intersections Based on the Mixed Traffic Agent Model of Autonomous-Manual Driving Connected Vehicles



You Ren, Shan Jiang, Guan Yan, Hongmei Shan, Huiying Lin, Zhilong Zhang, Liangzhe Wang, Xuesheng Zheng, and Jiaqi Song

Abstract Signalized intersection is an important node of urban traffic network, but also the most likely place in the urban traffic network to generate congestion, complex traffic conditions; and relatively high traffic flow will cause the reduction of urban traffic network operating efficiency, or even cause the paralysis of the urban traffic network. With the participation of pedestrians, the situation at the intersection will be more complex and diverse. In recent years, with the rapid development of automobile and related technology, manual driving connected vehicle (MDCV) and autonomous driving intelligent connected vehicles (ADICV) are gradually entering people's daily life. And ADICVs have coordinated and unified driving behaviors. In short, ADICVs have similar driving habits and can accelerate and decelerate at a systematic synchronization pace, which can help alleviate road traffic congestion and improve transportation network operation efficiency. Therefore, this paper uses the combination of ant algorithm and cellular automata to establish a pedestrian model, and constructs a traffic simulation model for the mixed traffic flow of the autonomous-manually driving vehicle agent at the intersection with/without pedestrian interference. Based on the assumption that the ADICVs have the coordinated driving behaviors, the paper analyzes the influence of green time ratio, mixed traffic ratio and average delay time on the traffic flow of MDCVs and ADICVs. Through the statistical analysis of the average velocity, traffic flow and other parameters, the mechanism and conclusion of the mixed traffic flow of MDCVs and ADICVs are obtained.

Keywords Mixed traffic flow of autonomous-manual driving connected vehicles · Ant algorithm · Cellular automata · Agent mixed model · Signalized intersection · Green time ratio · Mixing ratio · Average delay time

Y. Ren (✉) · S. Jiang · H. Shan · H. Lin · Z. Zhang · L. Wang · X. Zheng · J. Song
College of Traffic and Transportation, Jilin University, Changchun Jilin 130022, China

G. Yan
College of Mechanical Science and Engineering, Jilin University, Changchun Jilin 130022, China

1 Introduction

As early as 1920, the United States and other countries have begun to study autonomous vehicles [1]. So far, two complete autonomous driving classification systems have been established: SAE standard specified by the international society of automotive engineers, and the analysis system developed by the national high speed safety administration. In recent years, with the gradual popularization of 5 g network, intelligent connected vehicle has become an internationally recognized development direction with its advantages of more safety, efficiency, comfort and convenience [2].

In 2000, Sunnan Huang et al. proposed a autonomous vehicle driving system (AVDS), which avoids the need for vehicle-on-vehicle communication and can only use vehicle sensor information to improve the potential of capacity gain and smoother flow changing trend without reducing vehicle safety [3]. In 2003, Arnab Bose et al. analyzed the mixing phenomenon of semi-autonomous vehicles and manually driving vehicles and the basic diagram of its traffic characteristics. It was found that the mixed driving state of semi-autonomous or manual driving vehicles is compared with the driving state of purely manual vehicles, and the traffic flow of the former is greater at the same density [4]. Under parking and passing conditions, the average delay time of a semi-autonomous or manual vehicle mixed driving situation is shorter than that of a purely manual driving situation. However, under parking and passing conditions, there is no difference in the number of stationary vehicles between semi-autonomous or manual mixed driving and pure manual driving. Shih Ching Lo et al. proposed a cellular automata model for advanced vehicle control and safety system (AVCSS), the results show that the increase of autonomous driving vehicles will significantly improve the flow and speed of highway system [5].

In China, Xiaoping Qiu combined the Gipps vehicle following model with the cellular automata model to build a single-lane manual-autonomous vehicle mixed model in 2016. They pointed out that road capacity is positively correlated with the proportion of autonomous vehicles on the road [6]. In 2018, Wei Jianshe built a two-lane auto-manually driven vehicle mixed traffic flow model and found that the greater the number of auto-driving vehicles, the stronger the stability of the traffic flow. When the ratio of autonomous-manually driving vehicles reaches a certain ratio, the efficiency of the traffic system can be gained [7]. In 2019, Zhihong Yao and Rong Hu used the Full Velocity Differential Model (FVD) and the Cooperative Adaptive Cruise Control Model (CACC) to build a mixed traffic simulation model of manual driving vehicles and intelligent connected autonomous driving vehicles. This simulation analysis obtains the stability conditions of mixed traffic flow with different vehicle permeability [8].

In China, with the strong support of government, the research on autonomous driving by technology companies is also developing rapidly [9]. In 2013, Baidu established the Deep Learning Research Institute as a sign of entering the field of autonomous driving research. In 2017, Baidu proposed the Apollo software platform, and promoted the research of autonomous driving by sharing code in aspects such as

environment perception. Huawei is also making rapid progress in autonomous driving research. In 2017, Huawei and Vodafone cooperated. In 2018, Huawei cooperated with Audi again, and expanded the business development of driving automation and service digitalization, and launched 5G cars in 2020 [10].

In addition, as another element of the traffic system, pedestrians and vehicle often compete for the right of way at intersections in developing countries such as China, resulting in traffic delay and congestion, and even traffic accidents. Even at the signalized intersection, the difference of pedestrian's individual behavior and safety consciousness will seriously affect the traffic efficiency of motor vehicles.

In previous studies, the microscopic traffic flow simulation models established by scholars are more aimed at manually driving vehicles [11–13], and a few studies combine pedestrian models with vehicle models. However, the focus of the study is on the impact of vehicles on pedestrians [14–17]. The impact of pedestrians on motor vehicle traffic flow is rarely mentioned.

In summary, based on the assumption that autonomous driving intelligent connected vehicles (ADICVs) have coordinated and unified driving behaviors, this paper constructs a micro simulation model of mixed traffic flow of MDCVs and ADICVs, analyzes the traffic characteristics of MDCVs and ADICVs, and studies the characteristics of their mixed traffic flow at signalized intersection under pedestrian interference. The model can effectively reflect the traffic flow characteristics under the condition of mixed traffic of MDCVs and ADICVs, and the characteristics of their mixed traffic flow under pedestrian interference.

2 Construction of Mixed Traffic Model at Signalized Intersection Without Pedestrian Interference

Signalized intersection are the most complex places in urban traffic. Therefore, the actual situation in urban traffic should be considered as much as possible when constructing the simulation model of mixed traffic flow between MDCVs and ADICVs agent at signalized intersection:

- (1) In the construction of modern city, lanes can be divided into left turn, straight through and right turn. In this paper, in order to simplify the model as far as possible under the premise of simulating the actual road traffic situation, three lanes are set up. In short, they are left turn lane, straight lane and right turn lane.
- (2) The main intersections in the city are equipped with traffic signal lights, so the influence of signal lights should be considered in the model. And if the vehicle has the intention of lane changing, before reaching the intersection the vehicle must complete the corresponding lane changing action. Therefore, in the model, lanes can be divided into free lane changing area, forced lane changing area, and guide line area.

- (3) Due to the increased division of lanes, the original lane changing control of vehicle agent needs to be upgraded. Free lane changing module and forced lane changing module are added to the model.
- (4) In the simulation model, add the vehicle agent generation model to control the generation of vehicles in each lane.
- (5) Based on the assumption that autonomous driving intelligent connected vehicles (ADICVs) have coordinated and unified driving behaviors, vehicles not only need to pay attention to their position in the lane. According to the traffic regulations, when the signal light is red, vehicles need to stop instantly and wait when they reach the intersection. When the time left by the green light is not enough to support vehicles to pass through the intersection, vehicles need to slow down in advance.

In this model, the road characteristics at signalized intersections are shown in Table 1, and the road area division is shown in Fig. 1. When the vehicle agent is in the free lane changing area, or the vehicle agent perceives that the current driving situation does not meet its own needs, if the vehicle agent wants to pursue higher psychological expected speed or broader vision and space, the vehicle agent can freely change lanes in this area under the condition that its driving task can be completed.

In the forced lane changing area, which is close to the signalized intersections, the number of vehicle agents will be higher than the number of free lane changing area. If the current lane of the vehicle agent is different from the actual demand direction, the vehicle agent needs to change to the road consistent with its own driving direction through lane change control before reaching the guidance lane area.

Table 1 Road characteristics at signalized intersections

Direction of the lane	Turn left, go straight, turn right
Semaphore phase	Red light, green light
Lane division	Free lane changing area, forced lane changing area, guide line area

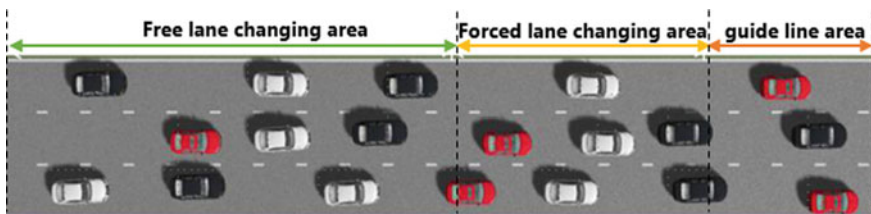


Fig. 1 Road schematic diagram of signalized intersections

2.1 Vehicle Agent Attribute Model of Signalized Intersection

In the simulation model of mixed traffic flow of autonomous or manual vehicle agent at three-lane signalized intersections, vehicle agent is the most basic and essential traffic factor. Therefore, when designing the vehicle agent attribute model, it should be able to express all the basic information of the vehicle agent when it travels. The basic information of the vehicle is defined as follows:

$$a(t) = \{v(t), L(t), T(t), R_n(t), R_o(t), L_v(t), v_{max}, a_{max}\} \quad (1)$$

where:

$v(t)$ —Velocity of vehicle agent at time T, unit is M/s;

$L(t)$ —The position of the vehicle agent on the lane at time T, and whether it is in the free lane changing area, the forced lane changing area or the guidance line area;

$T(t)$ —Type of vehicle agent, namely, self-driving vehicle or manually driven vehicle;

$R_n(t)$ —Lane number of the vehicle agent at time T, namely, it is in the left-turn special lane, the straight-ahead special lane or the right-turn special lane;

$R_o(t)$ —Target lane of vehicle agent;

L_v —The length of the vehicle agent;

v_{max} —The maximum allowable driving speed of vehicle agent in the current lane, unit m/s;

a_{max} —The maximum acceleration that the vehicle agent can generate when accelerating, unit m/s^2 .

2.2 The Control Model Generated by Vehicle Agent

In the three-lane signalized intersection autonomous or manual vehicle agent mixed traffic flow simulation model, the control model of vehicle agent is controlled by means of generating probability and entering probability. Moreover, the generation rate of straight-going vehicles, left-turning vehicles and right-turning vehicles meet the following formula:

$$P_z + P_l + P_y = 1 \quad (2)$$

$$P_c = P \leq P_d \quad (3)$$

where:

P_z —The probability of straight-going vehicle generation;

P_l —The probability of left-hand vehicle generation;

P_y —The probability of right-hand vehicle generation;

P_c —Generating function of vehicle;

P —Random probability;

P_d —Generation probability of vehicle agent.

2.3 Lane Changing Model of Vehicle Agent at Signalized Intersection

In the three-lane signalized intersections autonomous or manual vehicle agent mixed traffic flow simulation model constructed in this paper, the lane change behavior of vehicle agent in different lane changing areas is considered: when the vehicle agent is in the free lane changing area, the probability of vehicle lane changing motivation $P_{change} = 0.5$. When the vehicle agent is in the forced lane changing area, it will change lanes slowly according to its own demand. Do not change lanes when the vehicle agent reaches the guide lane area.

3 Construction of Mixed Traffic Model at Signalized Intersection with Pedestrian Interference

The signalized intersections with pedestrians is more complex than that without pedestrians. Considering the natural and actual behavior, the behaviors of pedestrians passing through the intersection are divided into the following categories:

- (1) According to relevant regulations, vehicles must yield to pedestrians when passing through intersections. Pedestrians, guided by traffic lights, choose the fastest and simplest way to cross the intersection, and pay little attention to the environment on the driveway. When they encounter obstacles at the intersection, pedestrians rarely choose to retreat, but try to pass through at one time by means of translation, etc.
- (2) When pedestrians pass through an intersection, they usually pass in groups with the characteristic of automatically forming a queue.
- (3) Very few pedestrians will choose to ignore the guidance of the traffic lights, and choose to make risky behavior through the intersection.

3.1 Construction of Pedestrian Update Model

The pedestrian model selects the combination of ant algorithm and cellular automata, and the pedestrian's walking direction is left to right and right to left. Based on the assumption that autonomous driving intelligent connected vehicles (ADICVs) have coordinated and unified driving behaviors, in order to simulate the natural and actual pedestrian behaviors in some cities in China, even when the traffic lights are red, pedestrians are allowed to pass through the intersection in this research model. In the model, the location of the pedestrian at the next moment is determined by the size of the pheromone value stored in the cell.

We calculated the pheromones produced by pedestrians passing to the right and left respectively, and the formula is as follows:

$$\tau_{i,j}^{all} = \tau_{i,j} + \tau_{i,j}^z \quad (4)$$

Right-handed pedestrian pheromone:

$$\tau_{i,j} = (1 - Rho) \times \tau_{i-1,j} - 1(k) + \Delta\tau_{i-1,j-1}(k). \quad (5)$$

Left-handed pedestrian pheromone:

$$\tau_{i,j}^z(k+1) = (1 - Rho)\tau_{i-1,j-1}^z(k) + \Delta\tau_{i,j}^z(k) \quad (6)$$

Substituting (5) and (6) into (4), can get:

$$t_{i,j}(k+1) = (1 - Rho) \times [t_{i,j}(k) + t_{i,j}^z(k)] + [\Delta t_{i,j}(k) + \Delta t_{i,j}^z(k)] \quad (7)$$

where, Rho is the attenuation value of pheromone.

When the vehicle signal light is red, the update rules of pedestrian position are as follows:

- (1) Pedestrian location initialization. Within each time step, pedestrians arrive randomly according to a certain probability. When pedestrians occupy the cell:

$$Position_{(i,j+1)}(k+1) = 1 \quad (8)$$

where: i represents the row where the pedestrian is and j represents the column where the pedestrian is.

- (2) The position at the next moment is determined by the cell. The pedestrian calculated the pheromone that could move at the next moment according to the method of ant cellular automata. And the cell with the highest pheromone value was selected as the cell at the next moment.
- (3) Update all pheromones once.

- (4) Update the location of pedestrians. When pedestrians reach the left or right edges, they leave the system; the rest of the pedestrians follow step 2 to select and update the pedestrian location.

3.2 Construction of the Interaction Model Between Pedestrians and Vehicles

Based on the assumption that ADICVs have coordinated and unified driving behaviors, because of the regulation effect of signal lights, both straight-going and left-turning vehicles pass according to the signal lights. When the red light is on, pedestrians and vehicles almost act as two separate groups without interfering with each other. When the color of signal light is about to change, pedestrians and vehicles will collide, causing interference.

- (1) When the signal light changes from green to red, the vehicles passing through the intersection will still continue to move forward. At this time, vehicles interfere with the normal passage of pedestrians. Pedestrians should not choose to pass through the intersection immediately according to the actual situation. When all vehicles have passed the intersection, pedestrians begin to move.
- (2) When the signal light changes from red to green, pedestrians who have not passed the intersection in time will still pass through the intersection. Vehicles that stop and wait at the intersection need to wait for all pedestrians to pass through the intersection before starting to pass through.

Based on the mixed traffic model of intersection without pedestrian signal, pedestrian passage is added to the model. The width of the crosswalk is 6 m and the length is 20 m. The cell size is specified as the floor area projected by the pedestrian's shape, and the cell space at the crosswalk of cellular space is 15×50 cells. The initial value of pheromone is set as $\tau_{i,j} = \tau_{i,j}^z = 1$, and the pheromone increment is $\Delta\tau_{i,j} = \Delta\tau_{i,j}^z = 1$. The rules set in this model are:

- (a) Pedestrians can only pass through the intersection from the pedestrian passageway. Pedestrians move in both directions simultaneously.
- (b) Vehicles and pedestrians cannot appear on the crosswalk at the same time. That is a vehicle cannot pass when there are pedestrians, and pedestrians cannot pass when there are vehicles.
- (c) When the red light is on, vehicle cannot pass and only pedestrians can pass through the intersection. When the green light is on, pedestrians can not pass, only vehicle through the intersection.

4 Research and Analysis of Mixed Traffic Flow at Signalized Intersection

The model above is used to simulate the traffic flow at signalized intersections and analysis the influence of green lights ratio and mixed traffic ratio on the traffic flow.

4.1 Simulation Process

Before starting the simulation, it is necessary to preset the basic parameters of the model. In order to restore the actual traffic conditions, we set the conversion period of the signal lamp to 100 s and the total length of test section is 1000 m. The length of free lane changing area is 700 m, the length for forced lane changing area is 250 m, and the length of the guide line area is 50 m. The maximum allowable velocity of the vehicle on the lane is 40 km/h, maximum acceleration of the vehicle when accelerating is 4 m/s² and unified vehicle length is 5 m. The model boundary adopts an open boundary, and the vehicle agent leaves the system when leaving the lane. The probability of a vehicle entering the system is 0.4.

The concept of average delay time is also introduced in the standard model, and the calculation formula is as follows:

$$\delta = \frac{T(1 - \lambda)^2}{2(1 - \varepsilon)} + \frac{\varepsilon^2}{2q(1 - \varepsilon)} - 0.65\left(\frac{T}{q}\right)^{\frac{1}{3}}\varepsilon^{(2+5\lambda)} \quad (9)$$

where:

δ —The average delay of each vehicle at a certain time on the lane, the unit is s;

Q —Traffic flow at one time (veh/s);

C —The maximum traffic capacity of the lane (veh/s).

ε —Road saturation, equal to traffic flow at a certain moment (veh/s)/the maximum traffic capacity of the lane (veh/s).

After the basic data preset is completed, the model simulation is carried out. The specific steps are as follows:

Step 1: Initialize the simulation program: determine the basic parameters, including the number of lanes, length of the guide lane, starting position of forced lane changing, total simulation time, signal cycle, etc.

Step 2: Generate vehicle agent, vehicle attributes (size, driving direction, velocity, acceleration, etc.);

Step 3: The vehicle agent changes lanes according to the demands: first, determine the area of the vehicle, and then make lane changing according to the corresponding lane changing rules;

Step 4: According to the type of vehicle, judge how to follow the vehicle. If it is an autonomous vehicle agent, iterate according to GIPPS rules. Otherwise, iterate according to the following rules of the manual driving vehicle agent;

Step 5: Determine whether the vehicle agent has completed the calculation. Step 6 is performed if the calculation is completed; otherwise, Step 3 is performed.

Step 6: Update all vehicle agents in the lane (velocity, position and other parameters);

Step 7: Determine whether the total simulation time is reached. If it is, Step 8 is performed; otherwise, return to step 2.

Step 8: Receipt simulation process data diagram;

Step 9: End the program.

4.2 Analysis of Simulation Results

In this simulation, we mainly study the influence of green time ratio and mixed traffic ratio on traffic flow, so the simulation was carried out twice. In the first simulation, by adjusting different green signal ratios, the influence of green signal ratios on the average vehicle density lane and average delay time at a certain time was studied. In the second simulation, by adjusting different mixing ratio, the influence of mixing ratio on the average speed and average delay time of vehicles at a certain moment was studied.

4.2.1 Analysis of the Impact of Green Time Ratio on Mixed Traffic Flow

Within a signal cycle, the influence of green time ratio on the average vehicle density is shown in Fig. 2. We can see that the impact of green time ratio on the average density of vehicles is obviously. With the gradual increase of green time ratio, the average density of vehicles gradually decreases, which indicates that the increases of the green time ratio are conducive to easing traffic congestion. However, when the green time ratio reach 0.8, the average density curve occurs obvious bending, and the slope of the curve decrease obviously. Along with the further increase of green time ratio, the average density curve almost tend to remain unchanged. This phenomenon shows that the set of the green time ratio are not the higher the better and too high green lights ratio can not reduce traffic congestion. The data shows that we should look for the most suitable green time ratio in the setting of urban signal lights. So as to improve the efficiency of regulation. In addition, as the proportion of autonomous vehicles increases, the overall average density of vehicles also decreases significantly.

The relationship between the average delay time of vehicles and the green time ratio is shown in Fig. 3. As can be seen from the picture, with the increase of the green time ratio, the average delay time of vehicles presents a very obvious downward

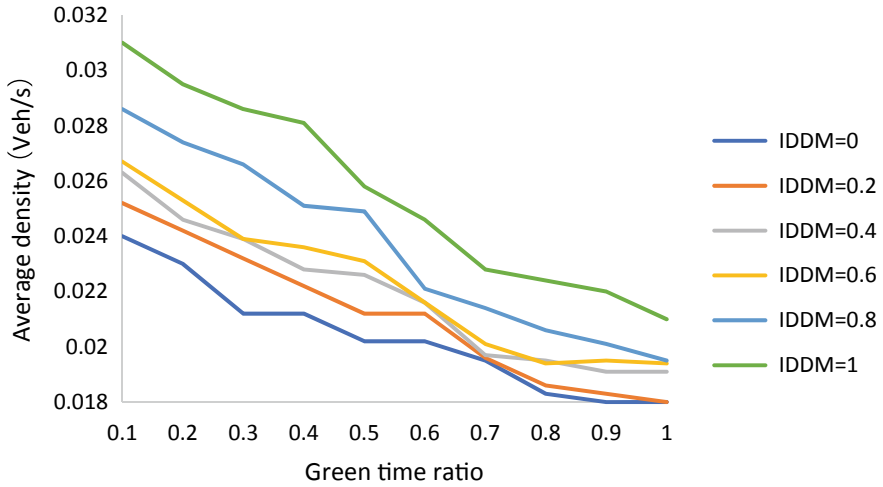


Fig. 2 Relationship between green time ratio and average vehicle density

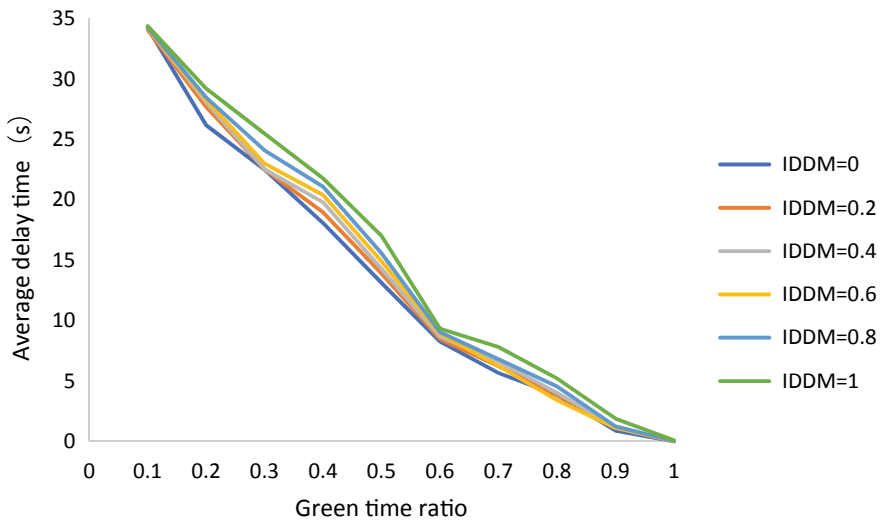


Fig. 3 Relationship between green time ratio and average delay time of vehicles

trend. This indicates that the increase of green time ratio have a positive effect on the decrease of average delay time. In addition, the mixing ratio of autonomous vehicle and manual vehicle has no obvious effect on the reduction of average delay time.

4.2.2 Analysis of the Influence of Mixed Traffic Ratio on Mixed Traffic Flow

The influence of mixed traffic ratio on the average velocity of vehicles is shown in Fig. 4. It can be seen that the proportion of mixed traffic has an obvious effect on the average velocity of vehicles. When the proportion of manual driving vehicles is large, the average velocity of vehicles is relatively slow. However, when the red light is in a signal cycle and the proportion of autonomous vehicles is large, the average speed of vehicles tends to decrease obviously. A possible reason for this phenomenon is that a manual vehicle can stay closer to the vehicle in front when the green light is on, so the acceleration of the vehicle must not be very large. While a self-driving vehicle will have a higher safety margin, allowing the vehicle to accelerate faster when the green light is on. When the red light is on, the autonomous vehicle will slow down in advance according to the driving behavior of the vehicle in front, resulting in the decrease of the average speed at this time. However, the manual driving vehicle may change lanes, which further aggravates traffic congestion and reduces the average speed of the entire lane.

The effect of mixed traffic ratio on average delay time is shown in Fig. 5. Similar to the average velocity, the larger the proportion of manual driving vehicles, the higher the average delay time. But with the increase of green light time, the average delay time presents a downward trend and can gradually keep stable. However, when the red light is on, the average delay time is significantly increased.

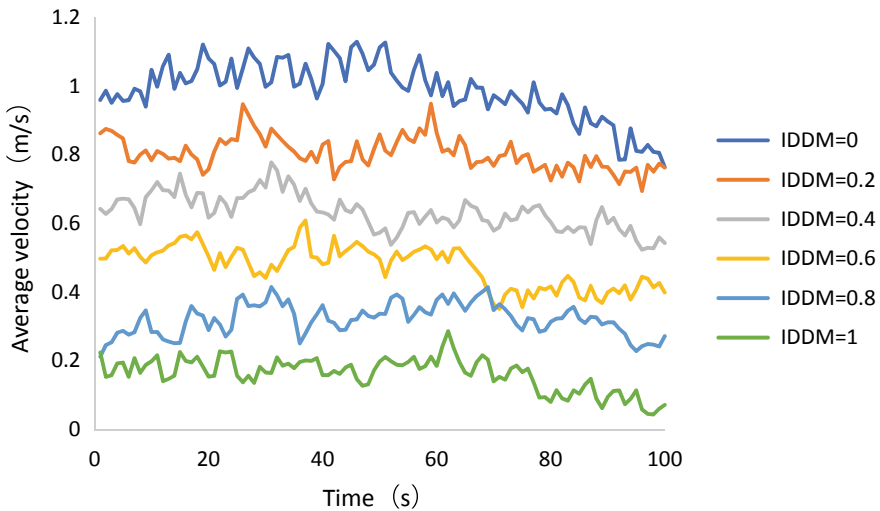


Fig. 4 Relationship between mixing ratio and vehicle average velocity

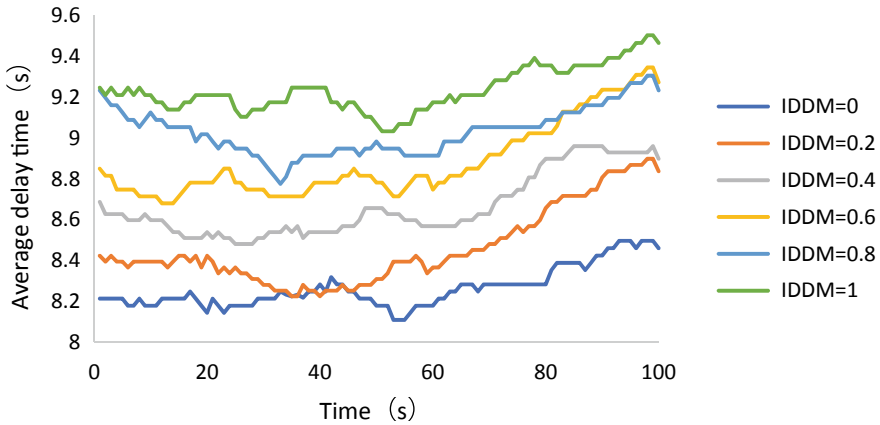


Fig. 5 Relationship between mixed traffic ratio and average delay time of vehicles

5 Conclusion

This paper analyzes the simulation results of the mixed traffic flow simulation model of autonomous and manual vehicle agent at signalized intersection based on the assumption that ADICVs have coordinated and unified driving behaviors. We can see that the green time ratio and mixed traffic ratio have a significant impact on traffic conditions. The green time ratio mainly affects the average vehicle density, and the proportion of mixed traffic has more influence on the average velocity and average delay time of vehicles. With the development of technology, more and more autonomous driving vehicles will enter people’s lives, which is very obvious to help alleviate urban congestion. However, in the current stage, more reasonable green time ratio can be adjusted to alleviate the pressure of urban traffic congestion.

Acknowledgements This work was supported by the Program of Ministry of Science and Technology “Vehicle–vehicle coupling mechanism and cooperative safety method , 2018YFB1600502”.

References

1. Yiyao X (2019) Development status and shape prediction of autonomous vehicles. China–Arab Sci Technol Forum (Chinese–English–Arabic) 2019(1):1–4, 105–109
2. Guanghui Z, Xiangyu L, Kai C (2019) Research on the development status of intelligent connected vehicles in China. Times Automobile 17:153–154
3. Huang S, Ren W, Chan SC Design and performance evaluation of mixed manual and automated control traffic. IEEE Trans Syst Man Cybern Part A (Syst Humans) 30(6):661–673
4. Bose A, Ioannou PA (2003) Analysis of traffic flow with mixed manual and semi-automated vehicles. IEEE Trans Intell Transp Syst 4(4):173–188

5. Lo SC, Hsu CH (2010) Cellular automata simulation for mixed manual and automated control traffic. *Math Comput Model* 51(7–8):1000–1007
6. Xiaoping Q, Lina M, Xiaoxia Z et al (2016) Research on manual-autonomous driving mixed traffic flow based on safe distance. *Transp Syst Eng Inf* 16(4):101–108
7. Xiujian W, Rongxin H, Hang S, Xiaoni Z, Hua C (2018) Game model and Simulation of two lane automatic manual driving vehicle mixed traffic flow. *Syst Eng* 36(11):97–104
8. Yao Z, Hu R et al (2019) Stability analysis and the fundamental diagram for mixed connected automated and human-driven vehicles. *Stat Mech Appl* 533:0378–4371
9. Jianyong M, Hongjun L (2017) Discussion on the development environment of China's intelligent connected automobiles. *Sci Technol Innov* 2017(02):1–2
10. Jianqiang W, Xin W (2017) Intelligent connected vehicle architecture and key technologies. *J Chang'an Univ (Social Sci Edn)* 19(06):18–25
11. Jingye L (2015) Research on agent-based traffic flow modeling. Nanjing University of Posts and Telecommunications
12. Jianhua Z, Yong P, Zhen C, Song L (2018) Single-lane continuous cellular automaton model based on IDDM-gipps hybrid strategy. *Chin Sci Technol Paper* 13(24):2760–2764
13. Juan L, Dayi Q, Cong L, Jinjin W, Xianghua X (2016) Research on vehicle lane changing behavior based on cellular automata. *Highway Transp Sci Technol* 33(11):140–145
14. Yingdong L, Huimin N, Jianqiang W, Shiwei L (2015) Research on traffic flow characteristics of crosswalk without signal. *Comput Eng Appl* 51(06):16–21
15. Jingjing L (2013) Pedestrian flow simulation analysis of signal control intersection crosswalk. Jilin University
16. Xiaolei D (2018) Analysis of pedestrian operation characteristics at signalized intersections and optimal design of crosswalks. Jilin University
17. Gaofei W (2018) Research on the influence of driver characteristics on traffic flow at signalized intersections. Chongqing Jiaotong University

Driving Velocity Tracking Error Analysis of Different Broadcast Methods Under Green Light Optimal Speed Advisory System



Shanshan Guo, Tao Zhang, and Yongsheng Liu

Abstract Green light optimal speed advisory (GLOSA) system provides the speed recommendations to drivers so that they can pass through successive intersections at the appropriate instants, aiming at saving time costs while maximizing energy efficiency. However, there is rare literature mentioning the impact of drivers' behaviors by different broadcast modes under GLOSA system. Combined with the direct "speed value" voice broadcast method, this paper compares the driving behavior differences under the effects of dashboard display of three different visual stimuli on driving behavior. The driving data of the driver when performing the optimal speed is collected from the real world, and the speed tracking error in the driving behavior is counted. Furthermore, the advantages and disadvantages regarding the three broadcast methods and voice broadcast method are illustrated with detailed analysis. The value and time recommendations of the "speed value" of the best speed consultation system are given at last. Thereby minimizes the error of the driver when performing the optimal driving speed.

Keywords GLOSA · Broadcast mode · Driver behavior · Speed tracking error

1 Introduction

There are lots of factors affecting the eco-driving energy consumption of the vehicle, which are mainly divided into the following categories: (1) powertrain type differences [1, 2]; (2) control strategy differences [3–6]; (3) driver differences [7, 8]. Compared with the powertrain type and control strategy, the influence of driver's

S. Guo · Y. Liu (✉)

Electromechanical and Vehicle Department, Weifang University, Shandong 261061, China

e-mail: lucy001478@163.com

T. Zhang

China North Vehicle Research Institute, Beijing 100072, China

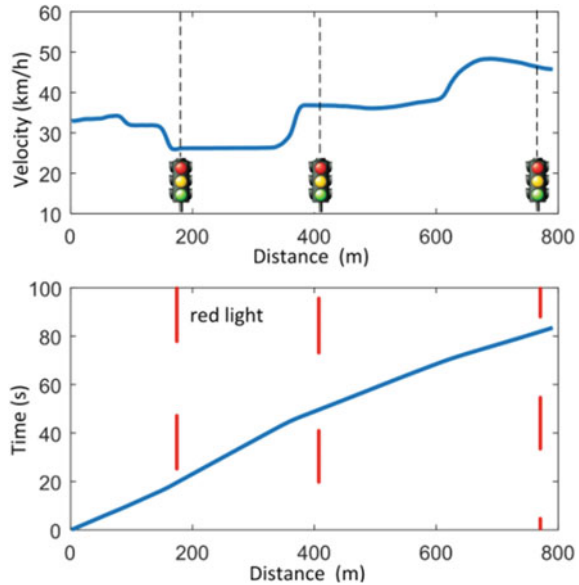
driving style on vehicle energy consumption is most important on urban roads. Unless the vehicle is completely autonomously controlling the driving, the drivers are ultimately responsible for ecological driving. The average fuel consumptions for different driving styles vary to the extent of 30% in urban environments and 17% on highways [7]. The driving style has a significant influence on the energy consumption of electric vehicles. Bingham et al. [8] found that the difference in electric vehicle energy consumptions between mild and aggressive driving styles was close to 30%.

The Green light optimal speed advisory (GLOSA) system refers to the in-vehicle intelligent system that rationally plans the optimal driving speed curve based on the state of the vehicle and the signal information provided by the urban intersection infrastructure [9]. The GLOSA application minimizes acceleration and braking time, and guides the driver to improve the economy of the vehicle while reducing time costs through following the speed suggested by the system. However, an important issue is a lack of current research on the way GLOSA broadcasts the advisory speed, which has a great impact on the driver's driving behavior. Obviously, the driver cannot stare at the screen frequently, on the other hand, the frequent replacement of the recommended speed will affect the driver's experience. Therefore, the appropriate broadcast method is especially important for the driver's acceptance. There is no exaggeration to say that the driver can drive the vehicle according to the optimal speed curve which depends on the way the system recommends the speed. So far, few people have been working to reduce the driver's driving error by studying the GLOSA broadcast method.

Literatures [10, 11] proposed that the GLOSA system provides the optimal speed of the road section for vehicles before the latest traffic signal. However, the article points out that the way in which speed consulting is implemented and how to provide signal scheduling information to vehicles is beyond the scope of their research. In [12], a dashboard pointer is used to display the GLOSA system, which saves 12% on average by taking driver error into account in advance. However, we have the reason to believe that simply relying on the screen to prompt the driver to change the speed is not a safe mode. In [13], it is simply mentioned that the eco-driving display interface is developed using a mobile phone to display the difference between the current vehicle speed and the recommended vehicle speed, but the research does not explain how the mobile phone software provided advice to the driver.

To the best of our knowledge, there are few reports on the driver behavior analysis of different broadcast methods under GLOSA system based on real-world eco-driving data. To fill this research gap, this paper starts from the planned optimal urban speed, and collects the driving error of the driver when artificially implementing the recommended speed. This paper analyzes the impact of different speed broadcast methods on driving behavior, investigates the driver's acceptance level, and then also points out the superior broadcast method of GLOSA.

Fig. 1 Optimal speed profile planned by the GLOSA system



2 Experimental Platform and Test Plan

2.1 Construct Advisory Velocity Curve

This paper use the dynamic programming algorithm and the model of the intersection signal mentioned in [5] to solve the optimal speed of multi-signal intersections with uncertain traffic signal timing. Since the focus of this paper is to collect the driving data of driver behavior when performing the recommended driving speed, the solution process of the optimal speed is not explained. The test road is a straight city road with the length of about 800 m and three small intersections. The optimal reference speed by DP optimization is shown in Fig. 1.

Unlike traditional solutions [14], updating the recommended speed in real time will upset the consistency of our subsequent data collection process. A fixed multi-segment recommendation speed that is planned in advance is used in this research.

2.2 Design of Test Platform

This paper develops a simple eco-driving broadcast system and the specific hardware solutions are shown in Fig. 2.

- (1) The inertial navigation (NAV982 GNSS) is assembled at the vehicle’s centroid position to obtain the vehicle’s travel speed and acceleration status information.

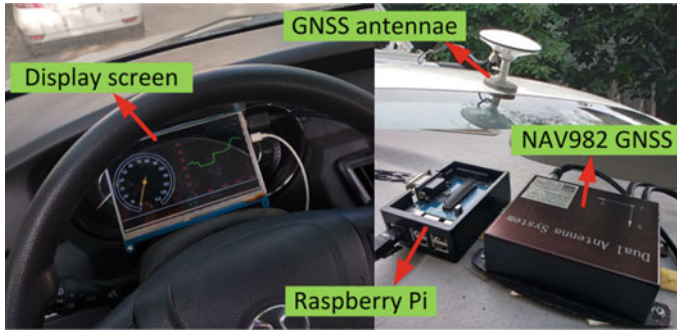


Fig. 2 Test platform of vehicle

Here the inertial navigation also provides vehicle GPS coordinate information for updating the real-time driving distance of vehicle.

- (2) By the robot operating system (ROS), the related built-in functions regarding drawing and voice broadcast can be called [15], which is convenient for the achievement of our proposed test platform. Finally, a screen that is close to the size of the dashboard is placed above the dashboard to provide advice to the driver with the recommended speed information.

2.3 Design of Different Broadcast Methods

Three different forms of dashboard screen display modes as shown in Fig. 3 is set up in this paper, which is able to provide related information in different visual forms to the driver.

Panel 1: The color block visual stimulus shows the interval of the suggested speed. The pointer indicates the actual vehicle speed change, as shown in Fig. 3a, the green square represents the recommended speed value, and the red square is the grid “5 km” range of the original data.

Panel 2: The digital visual stimulus shows the advisory speed and actual speed, as shown in Fig. 3b.

Panel 3: The curve visual stimulus shows the advisory speed. The future advisory speed curve of the continuous multi-stage optimization is displayed and the actual vehicle speed-distance curve is displayed in real time, as shown in Fig. 3c.

The broadcast system is also equipped to prompt the suggested “speed value” by the way of voice, which is convenient and more direct for the driver to obtain the recommended speed. There are some very important issues should be mentioned for adding the voice announcement function, i.e., how to choose the appropriate time point of the voice announcement and how to choose the appropriate broadcast speed value. Here we develop a “grid” average speed voice broadcast rule, as shown in Fig. 4:

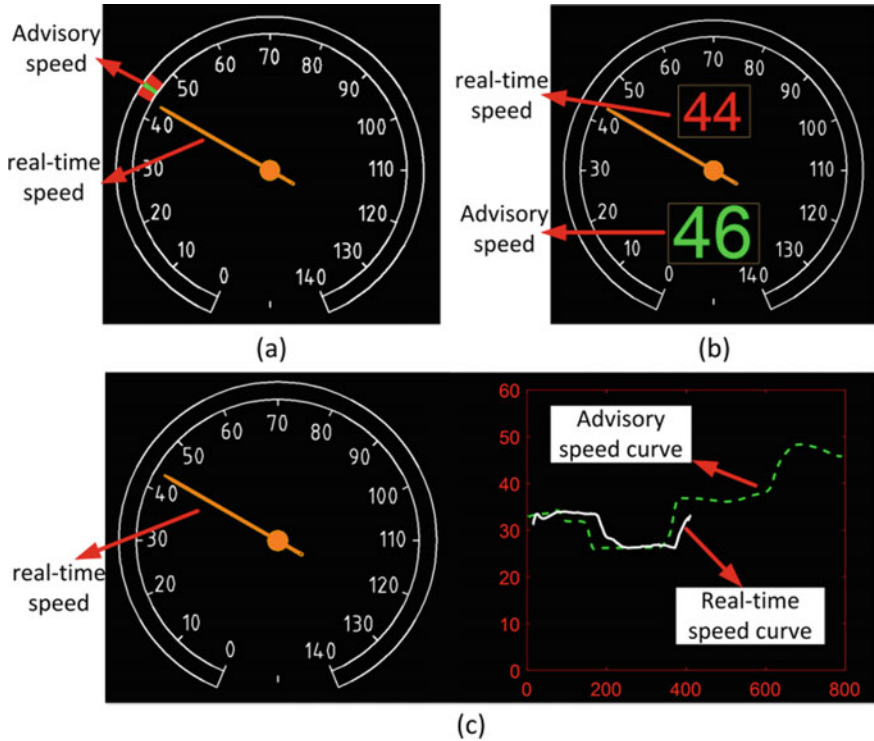


Fig. 3 Three different dashboard display modes

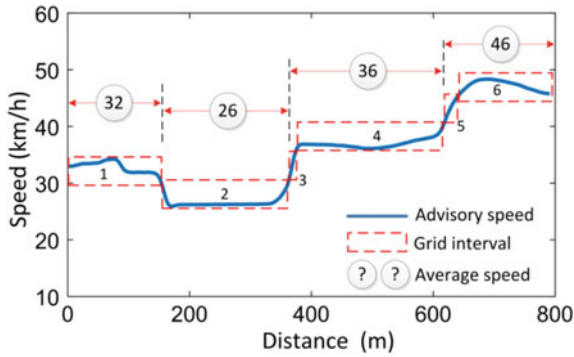


Fig. 4 The division of average speed curve

- (1) The speed interval grid (red dashed box) with a range of 5 km is applied to divide the advisory speed curve (blue line). The grid with time duration of less than 5 s is considered as the transition period (such as grids 3 and 5), and the rest are regarded as the uniform period with constant speed (such as grids 1, 2, 4, and 6).
- (2) The transition interval is used as the starting point of the uniform velocity interval, and the average velocity in the interval is broadcast. To approximate the effect of the actual vehicle display, we take the integers of the average speed values to be broadcast (32–26–36–46 km/h).
- (3) Perform a “speed value” voice prompt at the instant of speed change.

2.4 Real Vehicle Test and Data Collection

The following comparative tests are arranged:

Test 1: In the absence of a voice prompt, the advisory speed is displayed using three types of dashboards. The central axis of the green block in Panel 1 was the average velocity value. We used the numbers directly in Panel 2 to display the recommended average speed. Panel 3 displays the actual recommended speed profile (blue line in Fig. 4).

Test 2: In the case of voice prompts, the recommended speeds were displayed using three types of dashboards, and the same driving test was performed as in test 1.

Test 3: The time of the voice prompt was advanced, and the same test as in test 2 was performed.

We recruit ten drivers for this test, and the driver drives multiple times in each broadcast mode, collecting and filtering out at least 4 sets of valid data for each driver. The test roads corresponding to Fig. 1 is shown in Fig. 5.

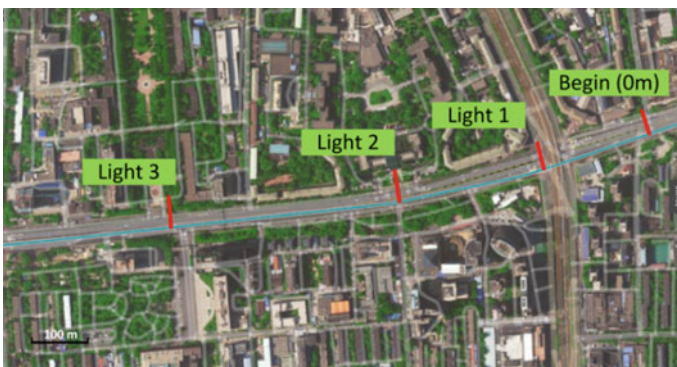


Fig. 5 The driving section

3 Results and Discussion

3.1 Impact of Dashboard Interface on Driving Behavior

Comparing the influence of different driving broadcast methods on driver’s driving error, this paper calculates the Mean Square Error (MSE) (1) between the actual driving data in the real world and the average advisory speed of the system broadcast prompt.

$$MSE = \sum_{s=1}^{800} [v_{real}(s) - v_{advisory}(s)]^2 / 800 \tag{1}$$

Figure 6 shows the MSE result of the driving data under different dashboard display modes without voice prompts. Figure 7 shows the driving data MSE under different screen broadcast modes when there is a voice prompt. All data statistics were fitted with a normal distribution and the expected value u of the response was calculated.

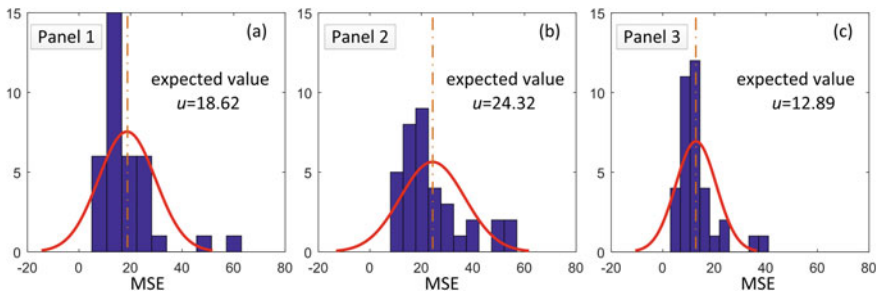


Fig. 6 MSE results without voice prompts

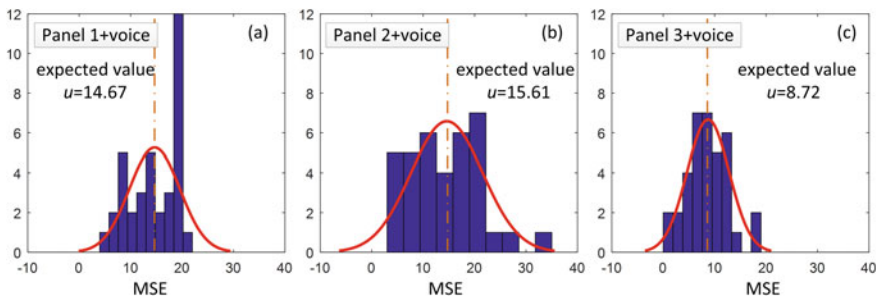


Fig. 7 MSE under s with voice prompts

Observe the MSE results of the driving data in the three broadcast modes in Figs. 6 and 7. The data shows that in any broadcast mode, it is difficult for the driver in the real world to track the recommended speed with a small error. The results of the two type tests show that the display mode 3 is more advantageous for reducing the speed deviation than the dashboard display mode 1 or 2 regardless of whether there is a voice prompt. This can be interpreted as a speed curve broadcast mode of panel 3, which can help the driver to observe the past speed tracking error of the vehicle. At the same time, the driver can preview the upcoming recommended speed value and be able to more intuitively understand the next stage of the action trend (acceleration, deceleration), preparing for action in advance and reducing reaction time.

For any dashboard broadcast modes, the error value in Fig. 7 is generally smaller than that of Fig. 6, which indicates that the broadcast mode with voice prompts can achieve the better performances. In Fig. 6, the recommended speed information is obtained only through the dashboard, resulting in a generally larger mean square error between the actual vehicle speed and the average recommended vehicle speed during manual driving. This can be explained by the fact that although the driver has been informed in advance the dashboard needs to be constantly observed to obtain the recommended speed information. However, it is still impossible to avoid the hysteresis caused by human observation, which causes the driver to fail to find the change of the recommended speed in time and generally results in the vehicle lags behind the suggested speed change moment. The prompts of the external voice help the driver to directly obtain the suggestion information and reduce the deviation during the manual driving process. Therefore, in order to obtain a better artificial driving effect, GLOSA needs to be equipped with the voice prompt function.

Further discussion of the driving difference with a voice prompt is depicted in Fig. 7. When the driver faces the dashboard display modes 1 and 2, the difference of driving performance is not obvious. The dashboard display mode 1 is slightly better than the display 2. Display modes 1 and 2 help the driver know the difference between the current actual speed and the recommended speed through different types of visual stimuli. Since the driver could not accurately control the speed of the vehicle, in mode 1, the driver tends to reach the expected speed when the speed pointer approaches the recommended interval of the speed. In mode 2, although the comparison of numerical values is intuitive, the driver does not continue observing the instrument panel after driving close to the recommended value of speed, and also performs free-floating driving. Therefore, the difference between the two broadcast modes is not particularly obvious.

3.2 Select Speed Value of the Broadcast

For the data samples corresponding to Fig. 7a and b, each driver had 8 sets of data (2 modes * 4 sets) and the proportion of time is calculated when the actual driving speed in each driving range is greater than the recommended speed. For example, during an 80 s in a whole journey, the actual vehicle speed is greater than

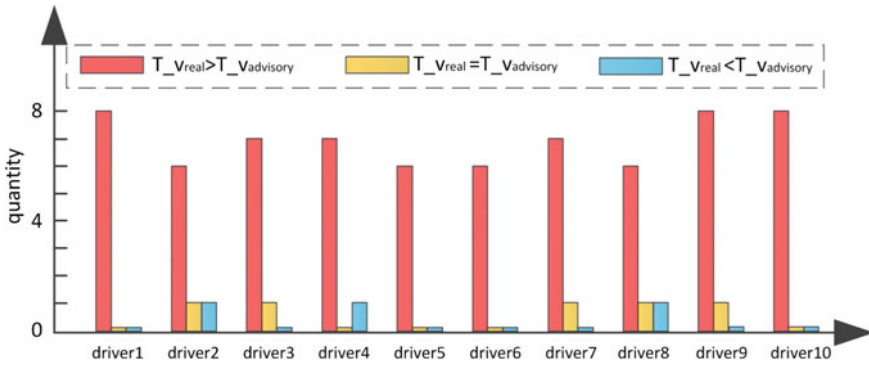


Fig. 8 The histogram of the number of drivers whose actual vehicle speed is greater than the recommended speed

the recommended vehicle speed for 45 s, and the actual vehicle speed is less than the recommended vehicle speed for 35 s, then the data is marked as $T_{v\cdot real} > T_{v\cdot advisory}$; If the difference between these two time lengths is within 5 s, the data is marked as $T_{v\cdot real} = T_{v\cdot advisory}$. The statistical results are shown in Fig. 8.

It can be seen from the statistical results in Fig. 8 that most of the actual driving speeds will be greater than the suggested speed guided by the system. To reduce the speed error caused by human factors, reducing the recommended speed value of the broadcast is necessary. As illustrated in Fig. 9, the difference between the mean value of the actual speed and the value of the advisory speed in the different vehicle speed ranges of the above samples is further counted. Obviously, the actual mean value of each speed interval is 0.66–0.83 km/h greater than the reported speed. Based on the above statistics, for the average speed of the GLOSA system planning, it is recommended to reduce the planned speed by 0.66–0.83 km/h firstly, then carry out the regular rounding broadcast. In the actual promotion and application, the driver data can be automatically collected by the in-vehicle system for analysis, thereby the value of the broadcast parameter is optimized.

3.3 Select Time Point of Broadcast

Figure 10 shows a typical driver acceleration and deceleration transient response curve. Under normal circumstances, the driver would not immediately execute the suggested speed while hearing the voice prompt. The speed changing lags behind the voice prompt due to the delay of the action caused by the driver’s observation and thinking.

The driver continues to respond to the start of the vehicle when he heard the voice broadcast recommendation “speed value”. It can be represented by the time taken from point A to point B in Fig. 10. According to the driving data of all 10 drivers in

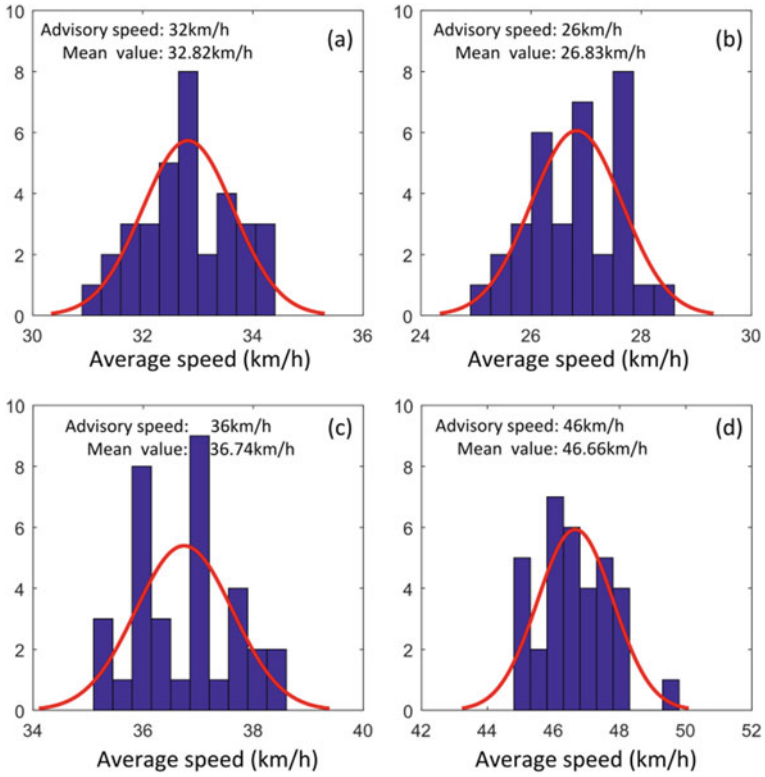


Fig. 9 Statistics of actual average speed values under different speed advisory

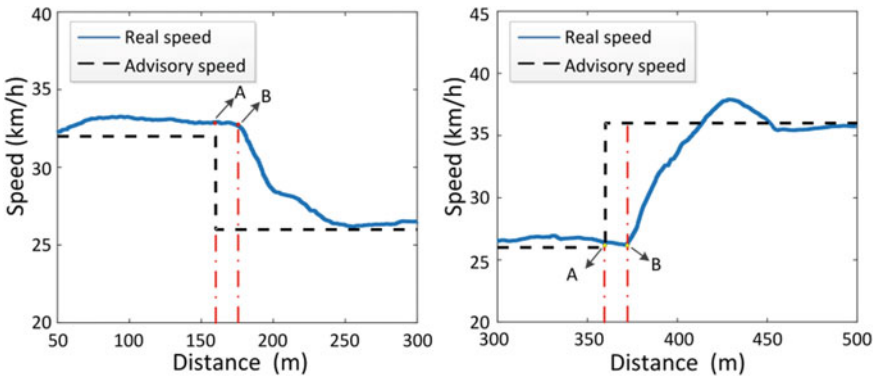
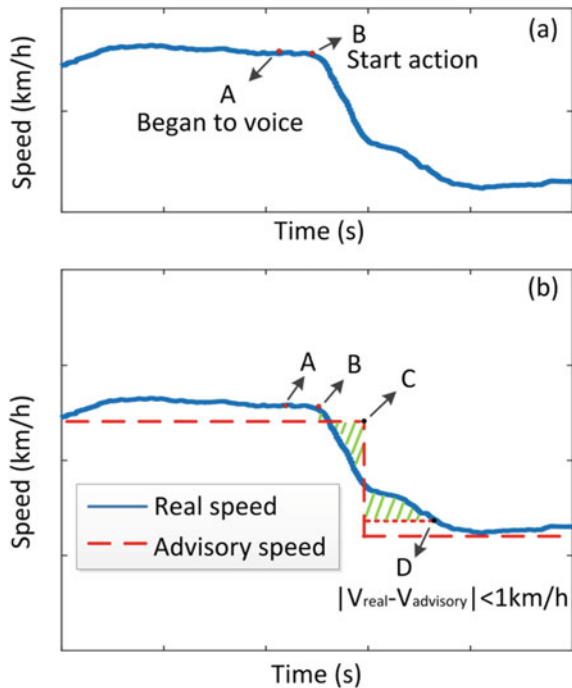


Fig. 10 The typical reaction rate curve

the three kinds of dashboard display modes, two sets of data are collected, and there are three voice changes cue points in each set of data. Therefore, the single driver reaction time data include 18 groups of data. The division error of the time point can be ignored. The average response time of each driver is 1.46 s, 1.51 s, 1.56 s, 1.58 s, 1.59 s, 1.62 s, 1.68 s, 1.68 s, 1.71 s, 1.75 s, respectively. In response to the reaction data of all drivers, the driver usually starts to control the vehicle speed after hearing the recommended “speed value” voice prompt, most of which are needed to be delayed by 1.5 s. Therefore, in the voice broadcast strategy of the GLSOA system, it is necessary to broadcast its recommended speed in advance to provide sufficient and appropriate timing advance for the driver to respond and perform actions.

In addition, the real vehicle speed cannot be abrupt and the suggested speed is a step change, which causes the actual speed of the vehicle to deviate from the desired average recommended speed during acceleration/deceleration process. How to determine the best speed broadcast time point and minimize driving the error are worthy of studying. In test 3, we did not perform a true repetitive test, but directly used the data collected in test 2 in advance. For the data in the combined broadcast mode of the dashboard display 1 and the voice prompt, as shown in Fig. 11a, a driver receives a voice prompt at point A and starts to operate until point B. In order to reduce the driving error caused by the behavior delay, the voice prompt can be changed in advance. As shown in Fig. 11b, point C is the recommended speed change point after the planning of the GLOSA system, and point D is the time when the actual

Fig. 11 Difference between the advisory speed and the actual response speed



vehicle speed approaches the recommended speed which difference is within range of 1 km/h for the first time. How to determine the starting point of the voice prompt (point A) can best offset the driver response delay time.

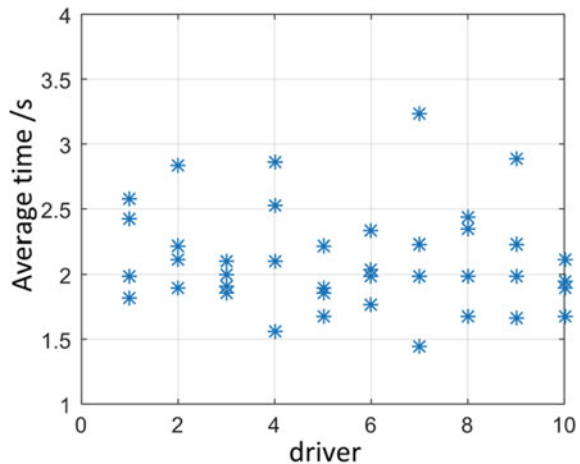
In the previous section, the statistics indicates that most of time intervals from A to B are more than 1.5 s. For a single driving data of a driver, the vehicle starts to perform at point B. The driving distance error $E_{B,D}$ from point B to point D is defined as the total errors between the actual speed and the advisory speed over the time horizon as follows:

$$E_{B,D} = \frac{1}{3} \sum_{i=1}^3 \sum_{t=t_B}^{t_D} [v_{real}(t) - v_{advisory}(t)] \tag{2}$$

where the constant 3 represents the number of advisory speed changes in a single driving data. At the same time, the statistical time from point B to point D is limited to 5 s regardless of the change in driving speed. The appropriate T_C time point determined so that $|E_{B,D}|$ is less than 0.1 km/h. The four groups of driving data for all 10 drivers are counted, and the results are shown in Fig. 12.

The time interval from T_B to T_C determined by the above statistical methods is mostly between 1.6 s and 2.8 s. Then the average time calculated is 2.13 s, which indicates that if the acceleration or deceleration of the vehicle is advanced by 2.13 s, the speed advance amount can be used to offset most of speed delay amount and reduces the driving speed error. Considering the above-mentioned driver behavior response delay time and action in advance time, the best broadcast time point A should start broadcasting at least 3.1 s before the recommended speed change point C.

Fig. 12 Time value statistical scatter plot from T_B to T_D



4 Conclusion

In this study, in view of the broadcast problems existing in the actual application of the GLOSA system, the influence of the system broadcast mode on driver behavior is comprehensively evaluated from the perspective of the true driving speed error. The research conclusions can be used to improve the real-world prediction application performance of GLOSA system:

- (1) In any broadcast mode, it is difficult for the driver to track the recommended speed with a small error in daily driving. The speed MSE under each kinds of broadcast mode satisfies the Gaussian distribution of positive, which has great research prospects in the future.
- (2) Among the three kinds of visual display broadcast methods of the GLOSA system, the recommended method that uses the color block style is a relatively popular and easy way for the driver to accept. Although the curve display method is more conducive to reducing the driving speed error, the similar interface will bring too much visual burden to the driver.
- (3) Voice prompt is more important than the dashboard display. In the broadcast system with voice prompt, the driving error is significantly reduced. The reason is that the driver receives the information more directly from the voice prompt, which can reduce the constant focus on the dashboard, that improves driving safety.
- (4) No matter which dashboard display mode is adopted, the actual driving speed of the driver is often higher than the recommended speed. Therefore, it is necessary to reduce the reported speed value. In practical applications, the driver data can be automatically collected and analyzed by the in-vehicle system, thereby optimizing the value of the broadcast parameters.
- (5) Considering the actual driver's reaction delay and the problem that the vehicle speed cannot be jumped, the voice prompt should be given in advance about 3.1 s at least. So the driver and the vehicle will have enough time to response.

Acknowledgements This work was supported in part by Ministry of Education Collaborative Education Project (201802105006, 201801245026), Shandong Provincial Higher Education Science and Technology Plan Project (J18kA012, J16LB10, J17KA150); Weifang Science and Technology Development Plan Project (Grant No.2019GX013).

References

1. Xiangdong H, Kegang Z, Yutao L et al (2005) Study on a new split type of HEV powertrains. *Int J Veh Des* 38(1):96–105
2. Westbrook MH (2001) The electric car: development and future of battery, hybrid and fuel-cell cars. Institution of Electrical Engineers
3. Yang H, Rakha H, Ala MV (2017) Eco-cooperative adaptive cruise control at signalized intersections considering queue effects. *IEEE Trans Intell Transp Syst* 18(6):1575–1585

4. Zhang T, Zou Y, Zhang X et al (2020) Predictive eco-driving application considering real-world traffic flow. *IEEE Access* (99):1–1. <https://doi.org/10.1109/ACCESS.2020.2991538>
5. Sun C, Shen X, Moura S (2018) Robust optimal eco-driving control with uncertain traffic signal timing. In: 2018 annual American control conference (ACC), pp 5548–5553. IEEE
6. Guo N, Shen J, Xiao R et al (2018) Energy management for plug-in hybrid electric vehicles considering optimal engine ON/OFF control and fast state-of-charge trajectory planning. *Energy* 163:457–474
7. Gonder J, Earleywine M, Sparks W (2012) Analyzing vehicle fuel saving opportunities through intelligent driver feedback. *SAE Int J Passenger Cars-Electron Electr Syst* 5(2012–01–0494):450–461
8. Bingham C, Walsh C, Carroll S (2012) Impact of driving characteristics on electric vehicle energy consumption and range. *IET Intell Transp Syst* 6(1):29–35
9. Typaldos P, Papamichail I, Papageorgiou M (2020) Minimization of fuel consumption for vehicle trajectories. *IEEE Trans Intell Transp Syst* 21(4):1716–1727
10. Seredynski M, Mazurczyk W, Khadraoui D (2013) Multi-segment green light optimal speed advisory. In: 2013 IEEE international symposium on parallel and distributed processing, workshops and ph.d. forum, pp 459–465. IEEE
11. Seredynski M, Dorransoro B, Khadraoui D (2013) Comparison of green light optimal speed advisory approaches. In: 16th international IEEE conference on intelligent transportation systems (ITSC 2013), pp 2187–2192. IEEE
12. Qi X, Wang P, Wu G et al (2018) Connected cooperative ecodriving system considering human driver error. *IEEE Trans Intell Transp Syst* 19(8):2721–2733
13. Dib W, Chasse A, Moulin P et al (2014) Optimal energy management for an electric vehicle in eco-driving applications. *Control Eng Pract* 29:299–307
14. Guo L, Gao B, Gao Y et al (2017) Optimal energy management for HEVs in eco-driving applications using bi-level MPC. *IEEE Trans Intell Transp Syst* 2017:1–10
15. Richardson M, Wallace S (2014) Make: getting started with Raspberry Pi electronic projects with the low-cost pocket-sized computer. Maker Media, Inc.

Mechanical Modeling and Analysis Based on Fine Drill



Kuikui Feng, Haodong Zhang, Qian Cheng, and Wuhong Wang

Abstract The main purpose of this paper is to study the change law of the contact condition between the fine drilling tool and the workpiece surface during cutting. Firstly, the structure of the drill tip of the fine drilling tool is analyzed, and the contact force of the drill tip is effectively decomposed. Then, on the basis of considering the cutting edge radius, the cutting mechanics model of each part of the drill tip was established according to the theory of slip line field. Finally, drilling experiments under different cutting conditions were carried out to solve the contact coefficient between the tool and the workpiece surface. The research shows that the high feed leads to greater ploughing effect in cutting, resulting in sharp change of contact force, which provides a research direction for tool tip design.

Keywords Fine drill · Main cutting edge · Chisel edge · Slip line field · Ploughing

1 Introduction

As an important part of aerospace engine, the injection plate is distributed with thousands of small holes. The size range of these holes belongs to the mesoscopic scale [1], and their geometric parameters play an important role in propellant injection performance and combustion stability in rocket engines. Therefore, it is of great

K. Feng · H. Zhang · Q. Cheng · W. Wang (✉)

Department of Industrial Engineering, Beijing Institute of Technology, Beijing 100081, PR China

e-mail: wangwh@bit.edu.cn

K. Feng

e-mail: fengkk1014@163.com

H. Zhang

e-mail: ZHDZYW@163.com

Q. Cheng

e-mail: billqiancheng@163.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

W. Wang et al. (eds.), *Green Connected Automated Transportation and Safety*,

Lecture Notes in Electrical Engineering 775,

https://doi.org/10.1007/978-981-16-5429-9_63

significance to establish the mechanical model and analyze its mechanical characteristics in the process of fine-hole cutting to improve the machining quality of injection plate and reduce the engine failure.

Fine drilling is essentially similar to conventional drilling and has many similarities in mechanical properties of cutting. However, the natural difference in the size magnitude also causes the mechanical properties to be different, which will be affected by the size effect [2], cutting edge radius [3], minimum cutting thickness [4] and other factors.

Abouridouane et al. [5] analyzed the scale effect in the cutting process of twist drill and established a three-dimensional FEM of ferrite-pearlite based C45 steel to study the influence law of cutting edge radius on chip shape. Lucca et al. [6] studied the relationship between plastic deformed layer, cutting edge radius and minimum cutting thickness in micro cutting, and found that the surface morphology of plastic deformed layer was independent of the minimum cutting thickness and only affected by the cutting edge radius. Malekian et al. [7] established the minimum chip thickness model based on the principle of minimum energy consumption and the infinite shear strain method, and then concluded that the cutting radius of the tool was positively correlated with the minimum chip thickness. Son et al. [8] established an ultra-precision cutting model and believed that the cutting edge radius and the friction coefficient should be considered to control the minimum chip thickness. Based on FEM, Vogler et al. [9] predicted that the minimum cutting thickness of pearlite and ferrite materials was 0.2 and 0.3 times of the radius of the blunt circle of the cutting edge, respectively. It can be seen that the cutting edge radius and the minimum cutting thickness are the important factors to determine the cutting quality of fine drill, and there is a close relationship between them. However, the cutting edge radius is dynamically changed under the action of force at all times in the cutting process. We need to carry on the mechanical analysis to the whole process of fine drilling cutting, in order to effectively monitor the tool state and predict the problems that will arise in fine drilling.

This paper focuses on the interaction between drill tip and borehole surface and its effect on material properties. In this paper, a mechanical model of fine drill based on time is established to analyze the contact force of each part of the drill tip, especially the change of the drilling force at the cutting edge when the cutting tool is worn. Through the analysis of the experimental data, the contact coefficient of the cutting deformation zone was solved, and the relationship between the tool wear and the contact force was evaluated.

2 Experimental Setup

The cutting force of the tool in the drilling process was measured experimentally to verify the mechanical model of fine drilling in this study. Figure 1 show the schematic diagram of the experimental device. We Prepared a 1Cr18Ni9Ti stainless steel sample, the size is $\varnothing 100 \times 4$ mm, placed it in the table of the Kistler 9275

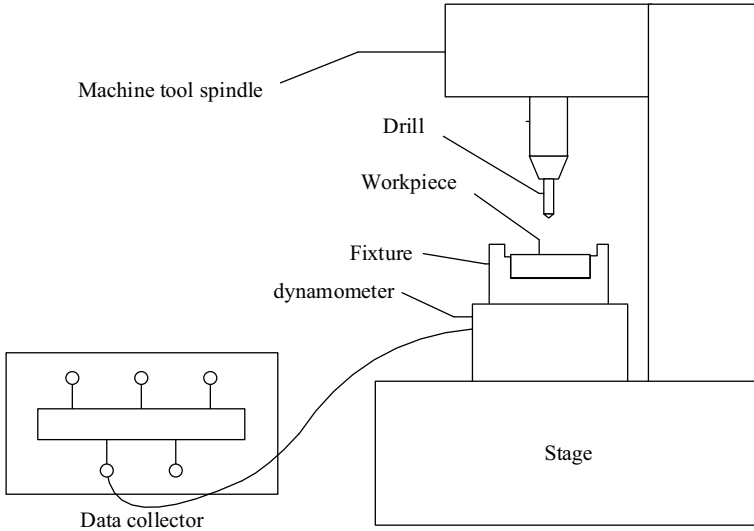


Fig. 1 Diagram of cutting experimental

Table 1 Parameters of tool

Cutting diameter	Chisel diameter	Blade angle
1 mm	0.3 mm	118°
Feed	Cutting speed	Rotate speed
5 $\mu\text{m}/\text{rev}$	18.8 m/min	6000 r/min

dynamometer in desktop on the dynamometer. The material of drill is carbide, its parameters are shown in Table 1.

3 Cutting Process Analysis

In the fine drilling process, the cutting part of the drill tip is composed of two parts, namely the main cutting edge and the chisel edge (Fig. 2). The main cutting edge is the intersecting line of the spiral groove and the flank, which is used for oblique cutting of the workpiece. The chisel edge is the intersection line between the rake face and the flank face, showing a slightly curved S-shape. Therefore, the cutting process of fine drilling is divided into five stages [10]. The 0– t_1 stage: The chisel cutting edge orthogonal cutting workpiece; The t_1 – t_2 stage: The main cutting edge oblique cutting material; The t_2 – t_3 stage: The drill tip is drilled into the workpiece completely, and the workpiece is acted on by two parts of cutting force. The t_3 – t_4 stage: Part of the chisel edge drill through the workpiece; The t_4 – t_5 stage: The chisel edge and part of the main cutting edge drill through the workpiece, and the workpiece

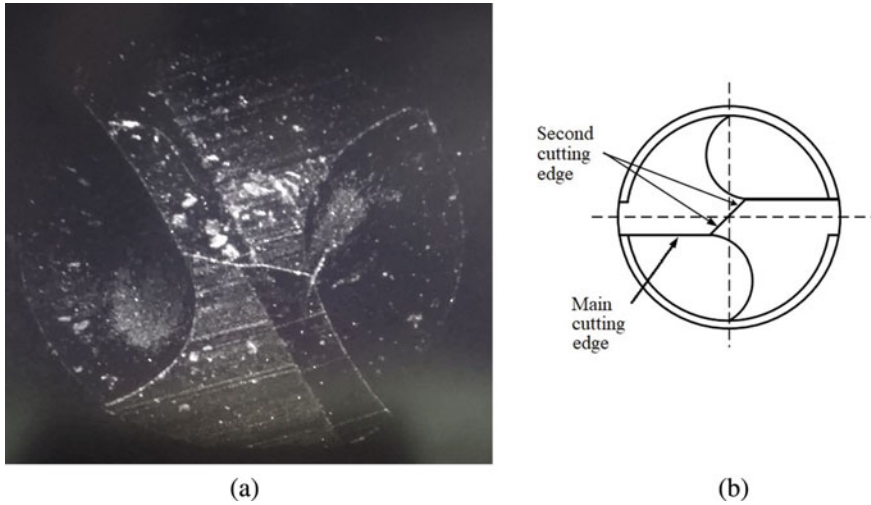


Fig. 2 Drill tip. **a** Entity, **b** structure

is only subjected to part of the cutting force of the main cutting edge until the cutting force is completely eliminated at the time of t_5 . So the cutting force model of five stages can be established, as follows in Eq. (1):

$$F_c(t) = \begin{cases} \frac{t}{t_1} F_s, & 0 < t < t_1 \\ F_s + \frac{t - t_1}{t_2 - t_1} F_m, & t_1 < t < t_2 \\ F_s + F_m, & t_2 < t < t_3 \\ \frac{t_4 - t}{t_4 - t_3} F_s + F_m, & t_3 < t < t_4 \\ F_s + \frac{t_5 - t}{t_5 - t_4} F_m, & t_4 < t < t_5 \end{cases} \quad (1)$$

where F_m and F_s are respectively the cutting forces of the main cutting edge and the chisel edge. Therefore, we can first establish the cutting force model of the two parts of the drill tip, and then solve the cutting force model based on time.

3.1 The Main Cutting Edge

The cutting process of the main cutting edge is oblique cutting (Fig. 3). The included Angle between the cutting direction and the chisel edge is the half peak angle φ , the thickness of the drilling core is r_w , and the distance between point a on the cutting

The model is divided into three plastic deformation zones [13]: the first shear plane region, ABCFGHIJK, the second deformation friction zone A'BC and the third deformation friction zone CSDEF. Surface CS and SD are edge contact surfaces, and their included angles with slip lines DE and CF are respectively ξ_1 and ξ_2 ; The central angles are respectively θ_b and θ_c ; The front Angle between AKJ and the horizontal plane is δ ; The shearing Angle between curve AB and the horizontal plane is ϕ_s ; The effective front Angle between CS and the vertical plane is γ_e ; The normal forward Angle between the rake face and the vertical plane is γ_n ; The radius of the drill tip is r_c .

The contact length between drill tip and workpiece are given as

$$\begin{cases} l_{CS} = 2r_c \sin\left(\frac{\pi}{4} + \frac{\gamma_n}{2} - \frac{\theta_c}{2}\right) \\ l_{SD} = 2r_c \sin \frac{\theta_c}{2} \end{cases} \quad (2)$$

The feed is f , so the thickness of the undeformed chip is given by

$$t_c = 0.5f \sqrt{\frac{r_i^2 - r_w^2}{r_i^2(1 + \cos^2 \varphi) - r_w^2}} \quad (3)$$

The contact length between the rake face and the chip:

$$l_{A'C} = \frac{t_c + \sqrt{2}(l_{CS} \cos \xi_2 + l_{SD} \sin \xi_1) \sin \delta - r_c(1 + \sin \gamma_e)}{\sin \phi_s} \quad (4)$$

The cutting force perpendicular to the cutting edge plane on the cutting element is

$$dF_m = k_s(k_{A'}l_{A'C} + k_C l_{CS} + k_S l_{SD}) \quad (5)$$

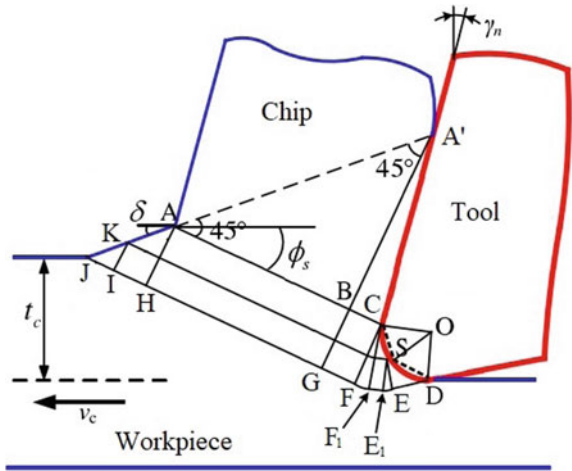
where k_s is the shear flow stress; $k_{A'}$, k_C and k_S are the contact coefficients between the tool and contact surface, respectively.

The axial cutting force F_m on the main cutting edge as follows:

$$F_m = (D - D_c)dF_m \cos \varphi \quad (6)$$

where D is the diameter of cutting tool; D_c is the diameter of chisel edge.

Fig. 5 Slip line field model of cutting element on second cutting edge



3.2 The Chisel Edge

Similarly, the chisel edge is equally divided into multiple cutting elements, and the cutting force acting on the workpiece is the superposition of the cutting force vectors of all the cutting elements. The orthogonal slip line field model of the chisel edge is shown in Fig. 5.

As in the model of the main cutting edge, the chisel edge is still divided into three plastic deformation zones. According to Eqs. (2)–(4), the contact length between the tool and the material is solved. Since the chisel edge is perpendicular to the cutting direction, the half peak Angle of the drill tip is $\varphi = 90^\circ$, and the thickness of unchipping t_c is $0.5f$. So $l_{A'C}$ is different. As the friction zone between the flank face and the workpiece have other zone-CFF1 and SEE1, ξ_1 and ξ_2 are reduced.

According to Eq. (5), We could calculate the cutting force of the cutting unit, then the axial cutting force on the chisel edge of the fine drill as follows:

$$F_s(t) = D_c d F_s \tag{7}$$

3.3 Cutting Force Experiment

The cutting heat generated by cutting will aggravate the tool wear and increase the radius of cutting edge, resulting in the undeformed chip thickness less than the minimum chip thickness and a large number of elastic deformation of the workpiece. The contact force is generated by ploughing and the tool contact coefficient will change at this time. Therefore, the cutting edge of the tool is actually dynamic, but in this example, it is assumed to be a constant value.

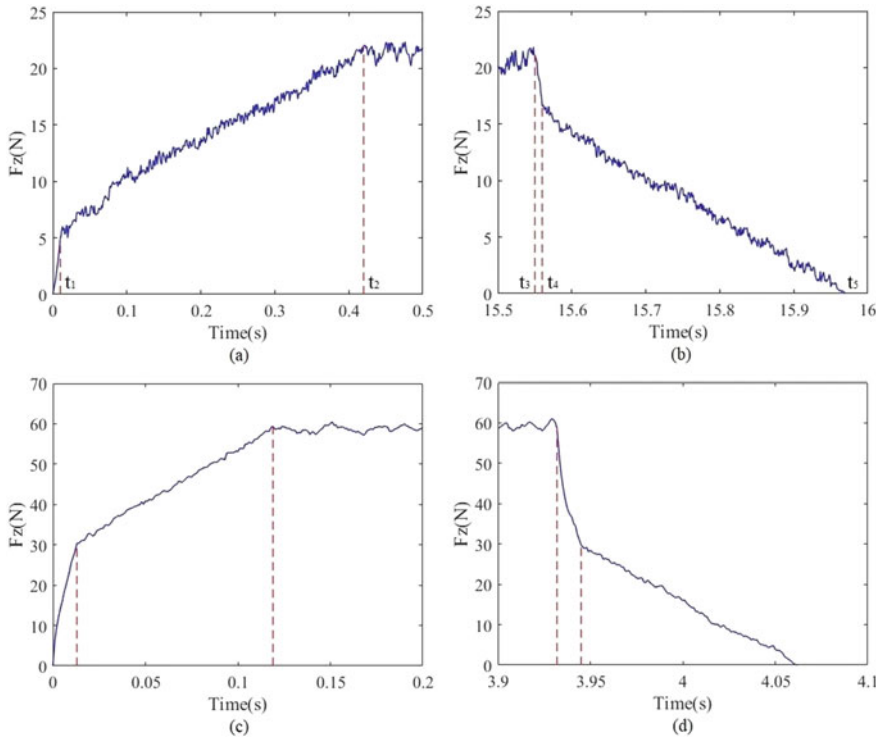


Fig. 6 The thrust value at low feed and high speed **a** $f = 5 \mu\text{m}/\text{rev}$ drilling in; **b** $f = 5 \mu\text{m}/\text{rev}$ drilling out; **c** $f = 20 \mu\text{m}/\text{rev}$ drilling in; **d** $f = 20 \mu\text{m}/\text{rev}$ drilling out

In order to calculate the cutting force of the tool in the cutting process, the unknown indentation half Angle and the contact coefficients need to be determined by micro-drilling experiments. We carried out cutting force experiments at constant rotational speed ($n = 6000 \text{ rpm}$) and different feed rates ($f = 5\text{--}20 \mu\text{m}/\text{rev}$). Figure 6 shows the thrust values for low and high feed cutting with an initial edge radius of $3 \mu\text{m}$. The chisel edge is drilled into the workpiece in a very short period of time, resulting in a very rapid increase in thrust, while the main cutting edge thrust increases slowly and remains stable after the bit is fully drilled. The thrust fluctuation value of the tool is large and the main cutting edge plays the leading role at low feed condition. It is related to the vibration in the cutting process. As the feed rate increases, the thrust value of the transverse cutting edge will become larger and larger, and even tends to exceed the main cutting edge, which is consistent with the description of Anish et al. [14]. It shows that with the increasing of tool temperature and wear, the contact conditions between the drill tip and the workpiece are changed, which affects the thrust measurement value.

Based on thrust measurements from fine drilling experiments under different conditions and the mechanical models of each area in Sect. 3, the contact coefficients

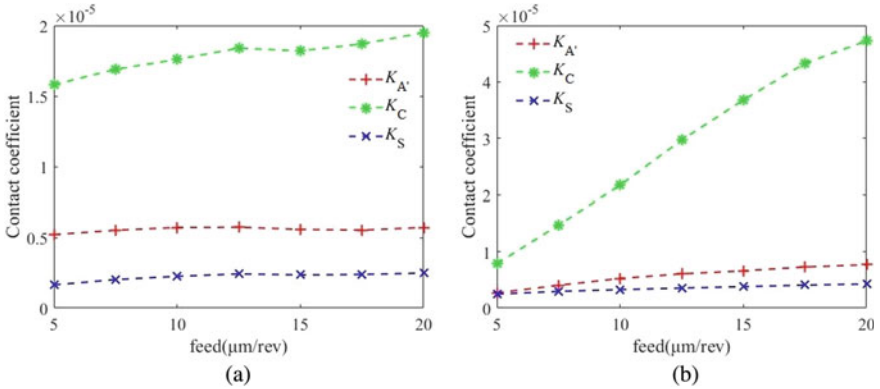


Fig. 7 Contact coefficient of tool-workpiece. a The main cutting edge, b the chisel edge

between the main cutting edge and the chisel edge of the drill tip and the contact surface can be determined by combining.

Figure 7 shows that no matter the main cutting edge or the chisel edge, the contact coefficient is much higher than the other two contact coefficients at different feed, which indicates that the resistance of the tool mainly concentrates on the intersection of the rake face and the flank face—the material shear area. With the increase of feed, the contact coefficients of main cutting edge and chisel edge trend. However, the thrust of chisel edge rise rapidly in a very short time, gathered in the energy. So the center part have serious ploughing phenomenon, which makes the chip more difficult to form. It is why the contact coefficient of chisel edge increases significantly faster. In this case, the tool wear is serious; the cutting edge radius r_c is increased; and the contact force is further improved. It also confirms the chisel edge thrust growth faster than the main cutting edge.

Therefore, when we design the structure of fine drill, we should appropriately reduce the diameter of the chisel edge to prevent chip accumulation on the basis of maintaining the tool stiffness.

4 Conclusions

- (1) Based on the consideration of the cutting edge radius, this paper analyzes the mechanical problem of the drill tip in the dry drilling process of stainless steel 1Cr18Ni9Ti sample, and puts forward an instantaneous mechanical model to predict the change law of the cutler-workpiece contact force and contact coefficient with the feed amount. The results show that the high feed increases the complexity of the contact conditions between the tool and the workpiece surface.

- (2) The contact coefficient at the intersection of the rake face and the flank face plays a leading role in the contact force of the whole tool. It is necessary to fully consider the local overheating caused by chip accumulation in the tool design, and reduce the width of chisel edge appropriately.

References

1. Zmijanovic V, Leger L, Depussay E, Sellam M, Chpoun A (2016) Experimental–numerical parametric investigation of a rocket nozzle secondary injection thrust vectoring. *J Propul Power* 32(1):196–213
2. Sabana A, Soumya G, Sankar MS, Rinku KM, Anandita S, Ramesh KS (2019) Study of cutting forces and surface integrity in micro drilling of a Ni-based super alloy. *J Manuf Process* 45:368–378
3. Abdellaoui L, Khelifi H, Sai WB, Hamdi H (2020) Tool nose radius effects in turning process. *Mach Sci Technol* 25(1):1–30
4. Aslantas K, Alatrushi LKH, Bedir F, Kaynak Y, Yilmaz N (2020) An experimental analysis of minimum chip thickness in micro-milling of two different titanium alloys. *Proc Inst Mech Eng Part B J Eng Manuf* 234(12):1486–1498
5. Abouridouane M, Klocke F, Lung D et al (2012) Size effects in micro drilling ferritic–pearlitic carbon steels. *Proc CIRP* 3:91–96
6. Lucca DA, Rhorer RL, Komanduri (1991) Energy dissipation in the ultra-precision machining of copper. *Annals CIRP* 40(1):69–72
7. Malekian M, Mostofa MG, Park SS et al (2011) Modeling of minimum uncut chip thickness in micro machining of aluminum. *J Mater Process Technol* 5:1–7
8. Son SM, Lim HS, Ahn JH (2005) Effects of the friction coefficient on the minimum cutting thickness in micro-cutting. *Int J Mach Tools Manuf* 45:529–535
9. Vogler MP, DeVor RE, Kapoor SG (2004) On the modeling and analysis of machining performance in micro-endmilling, part I: surface generation. *J Manuf Sci Eng* 126(4):685
10. Karpat Y, Karagüzel U, Bahtiyar O (2020) A thermo-mechanical model of drill margin-borehole surface interface contact conditions in dry drilling of thick CFRP laminates. *Int J Mach Tools Manuf* 154:103565
11. Chandrasekharan V, Kapoor SG, Devor RE (1995) A mechanistic model to predict the cutting forces in drilling: with application to fiber reinforced composite materials. *J Eng Ind* 117:559–570
12. Sambhav K, Tandon P, Kapoor SG, Dhande SG (2013) Mathematical modeling of cutting forces in microdrilling. *J Mech Sci Technol* 135(1):014501
13. Fang N (2003) Slip-line modeling of machining with a rounded-edge tool—Part I: new model and theory. *J Mech Phys Solids* 51(4):715–742
14. Paul A, Kapoor SG, DeVor RE (2005) A chisel edge model for arbitrary drill point geometry. *J Manuf Sci Eng* 127:23–32

Author Index

A

Ahmed, Tufail, 137

B

Burnett, Gary, 423

C

Cai, Guoqiang, 437
Cai, Mante, 723
Cai, Yan, 665
Cao, Honglin, 109
Chen, Chaozhou, 707
Chen, Deqi, 773
Chen, Di, 797
Chen, Guanliang, 723
Cheng, Guozhu, 597
Cheng, Qian, 411, 845
Chen, Ke, 219
Chen, Xizheng, 291
Chen, Yangzhou, 137
Chen, Yanyan, 759
Chen, Ye, 649
Chu, Guanquan, 479
Cui, Kaijun, 759

D

Deng, Hai, 163
Dong, Guoliang, 707

F

Fang, Xuezhao, 333

Feng, Kuikui, 411, 845
Feng, Qi, 109
Feng, Xiang, 665

G

Gao, Wei, 1
Gao, Yang, 163
Guo, Bin, 397
Guo, Hanying, 257
Guo, Jiawen, 291
Guo, Shanshan, 831
Guo, Wan-jiang, 563
Guo, Yingchun, 665

H

Han, Chao, 479
Han, Fengchun, 349
Han, Lifei, 245
Han, Lijin, 291
Han, Wan-li, 563
Hao, Siyuan, 523
He, Xuejing, 321
Huang, Zhen, 69

I

Ijaz, Muhammad, 137

J

Jamal, Arshad, 137
Jia, Hong, 789
Jiang, Jian-Feng, 219

Jiang, Lili, 377
 Jiang, Shan, 817
 Jiang, Xiaobei, 291, 411
 Jiao, Chengwu, 523
 Ji, Baoliang, 69
 Ji, Yanjie, 51

K

Kong, Weili, 533

L

Lan, Majing, 271
 Lei, Zhiyi, 549
 Liang, Jingqiang, 69
 Li, Chengbing, 177
 Li, Chunming, 509, 621
 Li, Dazhou, 1
 Li, Hao, 789
 Li, Jiahui, 523
 Li, Junzhuo, 397
 Li, Linghan, 797
 Li, Na, 349
 Lin, Huiying, 817
 Lin, Zihe, 363
 Lin, Zi He, 635
 Li, Qiliang, 163
 Li, Rong, 533
 Li, Shu-Bin, 195
 Lisheng, Jin, 163
 Li, Shurong, 773
 Liu, Bingjian, 423
 Liu, Dongwei, 51
 Liu, Fujia, 707
 Liu, Hao, 585
 Liu, Haoran, 321, 333
 Liu, Jiarui, 759
 Liu, Jun, 195, 607
 Liu, Ling, 607
 Liu, Pan, 309
 Liu, Si Yang, 635
 Liu, Wei, 69
 Liu, Xiaobing, 773
 Liu, Yang, 51
 Liu, Yongsheng, 831
 Liu, Yuanpeng, 457
 Liu, Yun-You, 219
 Liu, Zhenyu, 177
 Li, Wenyong, 397
 Li, Ying, 83
 Li, Yue, 447, 689
 Li, Zhixiao, 723

Lou, Xiaofeng, 745
 Lou, Yining, 231, 245
 Lu, Jiahuan, 585
 Lv, Zhiyong, 97

M

Ma, Guosheng, 723
 Ma, Jianming, 363
 Ma, Ke, 577
 Ma, Minglu, 377, 389
 Mao, Jian, 649
 Ma, Yanli, 219, 231
 Ma, Zhikuan, 549
 Mu, Shenglin, 83

N

Nie, Xinyue, 363
 Niu, Danhua, 621
 Niu, Zhongchen, 83
 Ni, Xiaoyang, 83

O

Ou, Chuan-jin, 447
 Ou, Chuanjin, 689

Q

Qiu, Shi, 275

R

Ran, Jin, 109, 275
 Ren, You, 817
 Rong, Jian, 363, 635
 Ruan, Chaoyu, 271
 Rui, Liang, 21

S

Sabino de Sousa, Jéssica Saturnino, 549
 Shan, Hongmei, 817
 Shao, Chunming, 509, 621
 Shen, Zhe, 163
 Shu, Chenglong, 621
 Shu, Peng, 533
 Song, Jiaqi, 817
 Sun, Chuanjiao, 543
 Sun, Junyou, 321
 Sun, Shichao, 37
 Sun, Xiaoxia, 301, 509, 621
 Sun, Xu, 423

T

Tang, Kaixuan, 83
 Tang, Rong, 585
 Tian, Jia-Jia, 219
 Tian, Jun, 301
 Tong, Ting, 491
 Tong, Xiaoping, 457

W

Wang, Bing, 321, 333
 Wang, Guobo, 389, 479
 Wang, Guozhu, 509, 621
 Wang, Hanlei, 149
 Wang, Hao, 577
 Wang, Liangzhe, 817
 Wang, Lin, 109, 275
 Wang, Liwei, 773
 Wang, Shaofan, 275
 Wang, Wuhong, 291, 411, 845
 Wang, Yi, 635
 Wang, Yigang, 163
 Wang, Yuqian, 675
 Wei, Panyi, 543
 Wei, Wei, 607
 Wei, Yiyang, 69
 Wei, Zhonghua, 275
 Wu, Keman, 523
 Wu, Peng, 149

X

Xia, Hai-ying, 789
 Xia, Haiying, 707
 Xiao, Dianliang, 675
 Xie, Wei-guo, 563
 Xing, Xiao-qing, 649
 Xiong, Ziyuan, 69
 Xu, Chengcheng, 309
 Xu, Donghui, 759

Y

Yang, Lining, 621
 Yang, Nan, 649
 Yan, Guan, 817
 Yang, Zhao, 585
 Yang, Zhicheng, 177
 Yan, Xuedong, 773
 Yao, Guangzheng, 759

Yao, Yuan, 149
 You, Jiangping, 491
 Yu, Fenghe, 245
 Yu, Wenlin, 291
 Yu, Xinyu, 797

Z

Zahid, Muhammad, 137
 Zeping, An, 21
 Zhang, Bo, 607
 Zhang, Hao, 447, 457, 689
 Zhang, Haodong, 411, 845
 Zhang, Hong, 745
 Zhang, Honghai, 585
 Zhang, Hongwei, 689
 Zhang, Jingyu, 309
 Zhang, Jinxi, 109
 Zhang, Liwen, 533
 Zhang, Miao, 437
 Zhang, Miaoyan, 51
 Zhang, Tao, 301, 509, 831
 Zhang, Wenya, 665
 Zhang, Xueli, 689
 Zhang, Ya-ping, 649
 Zhang, Ye, 437
 Zhang, Yu, 649
 Zhang, Yuhao, 257
 Zhang, Yumeng, 349
 Zhang, Zhilong, 817
 Zhao, Dan, 349
 Zhao, Haixia, 491
 Zhao, Jian-you, 563
 Zhao, Kuo, 321, 333
 Zhao, Nale, 523
 Zhao, Ning, 301
 Zhao, Yu, 97
 Zheng, Shan, 21
 Zheng, Xuesheng, 817
 Zhou, Chenjing, 363, 635
 Zhou, Chuanmiao, 597
 Zhou, Jin-dong, 121
 Zhou, Muqiu, 271
 Zhou, Siyuan, 423
 Zhou, Yuyang, 797
 Zhu, Huijie, 665
 Zhu, Jieyu, 231, 245
 Zhu, Yifan, 549
 Zhuo, Xingxuan, 149
 Zong, Chengqiang, 447, 689