

Smart Innovation, Systems and Technologies 127

Xiaoqing Zeng
Xiongyao Xie
Jian Sun
Limin Ma
Yinong Chen *Editors*



International Symposium for Intelligent Transportation and Smart City (ITASC) 2019 Proceedings

Branch of ISADS (The International
Symposium on Autonomous
Decentralized Systems)



 Springer

Smart Innovation, Systems and Technologies

Volume 127

Series Editors

Robert James Howlett, Bournemouth University and KES International,
Shoreham-by-sea, UK

Lakhmi C. Jain, Faculty of Engg and IT, CAI, University of Technology Sydney,
Broadway, NSW, Australia

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

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

More information about this series at <http://www.springer.com/series/8767>

Xiaoqing Zeng · Xiongyao Xie ·
Jian Sun · Limin Ma · Yinong Chen
Editors

International Symposium for Intelligent Transportation and Smart City (ITASC) 2019 Proceedings

Branch of ISADS (The International Symposium
on Autonomous Decentralized Systems)

 Springer

Editors

Xiaoqing Zeng
Tongji University
Shanghai, Shanghai, China

Jian Sun
Shanghai Jiao Tong University
Shanghai, Shanghai, China

Yinong Chen
School of Computing, Informatics
and Decision Systems Engineering
Arizona State University
Tempe, AZ, USA

Xiongyao Xie
Department of Geotechnical Engineering
Tongji University
Shanghai, Shanghai, China

Limin Ma
School of Environmental Science
and Engineering
Tongji University
Shanghai, Shanghai, China

ISSN 2190-3018 ISSN 2190-3026 (electronic)
Smart Innovation, Systems and Technologies
ISBN 978-981-13-7541-5 ISBN 978-981-13-7542-2 (eBook)
<https://doi.org/10.1007/978-981-13-7542-2>

Library of Congress Control Number: 2019936512

© Springer Nature Singapore Pte Ltd. 2019, corrected publication 2019

This work is subject to copyright. All rights are reserved 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

Preface

The conference of ITASC (International Conference on Intelligent Transportation and Smart City) encompasses the topics of intelligent transportation, information technology, construction civil, and smart city technology. It aims to build a platform for communication and corporation. The 1st ITASC was held in March 2013 in Mexico, which started as the branch of the International Symposium on Autonomous Decentralized System (ISADS). Main technologies discussed at 1st ITASC have been utilized in rail transit in both China and Japan. Since then, the ITASC conference has been held every two years, and the latest ITASC in 2017 attracted over 50 specialists and nearly 600 people from both universities and corporations.

ITASC will be held in May 8–10, 2019, in Shanghai. The conference will provide a good chance of discussing topics on the latest developments in transportation, civil engineering, information technology, and environmental engineering. It will mainly focus on the theories and applications of modern technologies of intelligent transportation and smart city. It is hopeful to contribute harmonious life through combining them together theoretically. During the conference, we can spend pretty much time on discussing how to apply new techniques and products into transportation, municipal construction, and information control. Thus, the progress of the technology can be promoted in these areas.

In this proceeding, we chose 28 papers from more than 60 submitted articles, which cover several professional fields, such as transportation, civil, and environmental engineering. Quite a few of these articles chose specific perspective to describe new ideas of crossing different professional fields, which will inspire readers greatly. With the prosperous urban development in China, there are urgent needs and huge application prospects for the intelligent transportation and smart city, which actually have already been successfully implemented in some developed areas.

We believe there will be about 500–700 participants. We hope all those interested in improving our urban living qualities could attend our conference and share your wonderful discoveries and experiences, whether you are from universities,

companies, or governments. Hence, we give our sincere gratitude to all the sponsors, press, print, and electronic media for their excellent coverage of this conference.

January 2019

Xiaoqing Zeng

Contents

Real Time Gesture (Fall) Recognition of Traffic Video Based on Multi-resolution Human Skeleton Analysis	1
Xinchen Xu, Xiaoqing Zeng, Yizeng Wang, and Qipeng Xiong	
Determining Acceleration Lane Length on Expressway Weaving Area Using Microscopic Traffic Simulation	13
Jing Luo, Xiaoqing Zeng, Yinong Chen, and Daniel(Jian) Sun	
Study on Optimal Model of Traffic Signal Control at Oversaturated Intersection	28
Chaoyang Wu, Xiaoqing Zeng, Jifei Zhan, and Qipeng Xiong	
Application of Intelligent Transportation System in Intelligent Network Environment	39
Qiulan Wang and Sayi Wang	
Stochastic Traffic-Assignment with Multi-modes Based on Bounded Rationality	49
Zhi Zuo, Xiaofeng Pan, Lixiao Wang, and Tao Feng	
A Map Matching Algorithm Combining Twice Gridding and Weighting Factors Methods	63
Ketu Cao, Lixiao Wang, Zhi Zuo, and Xiaohui Sun	
A Study on the Decision-Making Heterogeneity of Parking Mode Choice	74
X. H. Li, L. X. Wang, X. H. Sun, and Z. Zuo	
An Improved Vehicle Detection and Tracking Model	84
Libin Hu, Zhongtao Li, Hao Xu, and Bei Fang	
Determination of Best Foaming Parameters of Modified Asphalt Using Response Surface Analysis Method	94
Wen Xiaobo, Wu Hao, Zhang Peng, and Ding Fan	

Adaptive Fuzzy Logic Traffic Signal Control Based on Cuckoo Search Algorithm	107
Suhua Wu, Yunrui Bi, Gang Wang, Yan Ma, Mengdan Lu, and Kui Xu	
Research on Road Traffic Signal Timing Method Based on Picture Self-learning	118
Guohua Zhu and Chi Zhang	
Review of Ramp Metering Methodologies for Urban Expressway	127
Songxue Gai, Xiaoqing Zeng, Chaoyang Wu, and Jifei Zhan	
Research on the Application of TOD Theory—Taking Huijin Road Station of Shanghai Rail Transit Line 17 as an Example	138
Jicheng Huang, Xiaoqing Zeng, Yizeng Wang, Yining Chen, Nixuan Ye, and Zhongzheng Ma	
Data-Driven Safety Model on Urban Rail Transit Signal System	145
Yujia Chen, Xiaoqing Zeng, and Tengfei Yuan	
Simulation Data Generating Algorithm for Railway Turnout Fault Diagnosis in Big Data Maintenance Management System	155
Ke Cui, Maojie Tang, and Dongxiu Ou	
Research on the Warning Threshold of Rail Transit Passenger Flow by Big Data	167
Tengfei Yuan, Xiaoqing Zeng, Qipeng Xiong, and Chaoyang Wu	
Optimization of Subway Departure Timetable by Genetic Algorithm	177
Junxiang He, Xiaoqing Zeng, Peiran Ying, Xinchen Xu, and Yizeng Wang	
Review of Safety Assessment System on Urban Rail Transit	186
Songxue Gai and Xiaoqing Zeng	
Research on Application of PC Technology in Shanghai Metro Line 17	197
Yizeng Wang, Xiaoqing Zeng, Nixuan Ye, Yining Chen, and Zhongzheng Ma	
Research on Test of Shear Strength of Aggregates Based on Gradation and Particle Shape	206
Chang-xuan He and Yan-feng Bai	
The Seismic Analysis of the Structure of Hydraulic Tunnels	214
Gang Wang	
Analysis on the Effect of Excavation by Sections for Large Foundation Pit Without Horizontal Struts	222
Dongqing Nie	

Research on Robot Education Curriculum System and Settings in Internet of Things Major 232
Ying Guo, Lianzhen Zheng, and Yinong Chen

Research on Expressway Network Tolling Platform Technology Based on GIS+BIM 241
Li-xia Bao and Yuan-qing Wang

Commercial Vehicle Dynamic Third Party Safety Supervision of Based on GPS/Beidou Technologies 251
Guangyue Nian, Jinping Li, Daniel(Jian) Sun, and Xinhua Li

Research on Application of BIM Technology in Municipal Engineering Construction 265
Xiaoqing Zeng, Qipeng Xiong, Yizeng Wang, Xinchun Xu, and Liqun Liu

Study on BIM Technology Application in the Whole Life Cycle of the Utility Tunnel 277
Chuanpeng Hu and Shilang Zhang

A Review of Air Quality Monitoring System Based on Crowdsensing 286
Min Huang and Xia Wu

Correction to: An Improved Vehicle Detection and Tracking Model C1
Libin Hu, Zhongtao Li, Hao Xu, and Bei Fang

Author Index 297



Real Time Gesture (Fall) Recognition of Traffic Video Based on Multi-resolution Human Skeleton Analysis

Xinchen Xu¹, Xiaoqing Zeng¹, Yizeng Wang^{2(✉)},
and Qipeng Xiong³

¹ The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, No. 4800 Cao'an Road, Shanghai 201804, China

² Shanghai University, Shanghai, China

1102790744@qq.com

³ Shanghai FR Traffic Technology Limited Corporation, Shanghai, China

Abstract. The objective of this study was to detect abnormal behavior event (fall) by analyzing the existing monitoring video for ensuring the safety of rail transit platform passengers. Here, the key point coordinates and limb joints of human body are obtained based on PAFs (part affinity fields). Then feature information determined the fall is extracted based on the neck key point tracking algorithm. The feature for judging fall, namely the angle between leg and horizontal plane, is proposed in this study. The experimental data is based on the simulated fall videos took from Leqiao station of Soochow 1st metro line in Jiangsu province. Our results show that, the single picture and video are tested respectively, and it is found that the timing information is more fault-tolerant and more accurate in identifying falls, yet is more complex and more difficult to implement. What's more, because of the lack of sufficient fall videos, the results analysis based on the proposed algorithm remains much room for improvements. Also, more useful and efficient detection characteristics can be taken into account in the future.

Keywords: Gesture recognition · Rail transit · Video surveillance · Fall detection

1 Introduction

Falls are rare, while not timely detection of falls poses a great threat to personal safety and health. In public area, it may trigger panic, crowd gathering, fleeing stampede and other larger mass incidents. With the rapid development of computer vision, pose estimation is attracting more and more attention. Fall recognition, as a special case of human abnormal posture, has become a research hotspot.

Some methods used in fall recognition to date include two methods based on non-visual and visual. Non-visual detection methods mainly rely on wearable sensing device to detect acceleration signals when people fall. However, wearing those sensors on the body brings certain inconvenience to the action. This method is mostly used for the supervision of the elderly or the patient, and is not suitable in daily life.

Now the mainstream approach is based on visual information such as video stream analysis based on depth camera, multi-camera and single fixed vision camera. With its depth information, the depth camera has more information and better effect than ordinary cameras, but its cost is higher and generality is not high. Multi-camera and depth cameras are very complex and costly, not suitable for those daily scenes. Monocular cameras are relatively simple and inexpensive to obtain. This article uses surveillance video from the single fixed-view camera. It can make full use of existing resources, and the results is expected to be applied to the reality.

Machine learning has been applied to more and more human pose recognition technologies. Tompson, Jain and others use the hybrid structure of depth convolution neural network and Markov random domain to recognize human pose [1]. Aslan, Sengur and others propose a fall detection system based on depth camera, where shape and support vector machine are used to distinguish falls from other daily behaviors [2]. The core technology in human behavior recognition is divided into three steps: human segmentation, feature extraction and representation, activity detection and classification. At present, the related methods of gesture recognition are mainly classified according to the models established. Non-model-based technology mainly extract features, and many body modeling are still 2D technology, which can obtain posture information more easily after labeling the human part [3].

PAFs (part affinity fields) [4] uses a non-parametric representation [4], learning to associate body parts with individuals in the image. It uses a bottom-up approach, which first regresses to the key points of everyone, and then divides the key points so that the key points can be assigned to everyone. The greedy parsing algorithm is sufficient to produce high-quality body posture analysis, even if the number of people in the image increases, the efficiency will remain efficient, which has a good application prospect for the scene of the mass transit platform. The real-time algorithm detects the two-dimensional pose of multiple people in the image, ensuring the timeliness of security monitoring and facilitating rapid rescue in emergency situations.

There are five main features of fall detection: aspect ratio (AR) [5], AR change, fall angle [6], center speed [6], and head speed [5]. These features all make full use of the posture difference between normal human and falling human. However, there are still few fall recognitions combined with the bone information of the human body. In order to identify the passenger's fall on the platform of rail transit, this paper uses daily surveillance video captured by the existing platform monitoring cameras. It is based on the current advanced bone extraction algorithm—PAFs (part affinity fields) [4], and analysis innovatively from the passenger's leg angle characteristics, through real-time calculation of the traveler's two thighs angle value and timing information.

2 Human Key Points Acquisition Based on Multi-resolution Part Affinity Fields Model

In this paper, the extraction of human key points and the connection of limbs and trunks are based on PAFs (part affinity fields) [4]. This method uses frame-by-frame processing to extract human skeleton without human detector in real time. Because of the large depth of rail transit platform, there may be a different number of available key point information in the surveillance video.

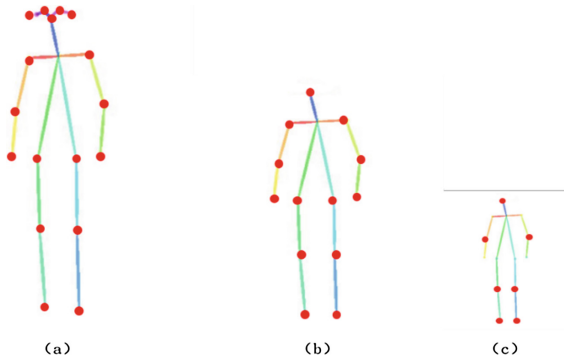


Fig. 1. Calibration of human key points in different depths of field

Therefore, a pedestrian attitude database of rail transit is established from video images, and different number of key points are calibrated for different depth of field pedestrian instances. As shown in Fig. 1(a), (b), (c) correspond to the calibration of key points of human body in close, middle and long-range respectively, which can effectively distinguish the information of people in different depth of field. The original model is

retrained using the database, and a new model with better recognition ability for rail transit scene is finally obtained. This paper uses this new model.

3 Fall Recognition Based on Multiresolution Part Affinity Fields Model

3.1 Leg–Horizontal Plane Angle Characteristics

Decision Rule. When a person falls, he will be tilted or even lie down. At this time, his legs will have a certain tilt angle, while the normal standing people's legs are almost upright, and for walking people, the angle will not be too large. In this paper, the angle of the two key points connecting the hip and knee, that is, the tilt angle of the thigh, is taken as the judgment condition of the fall, and the angles of the left and right

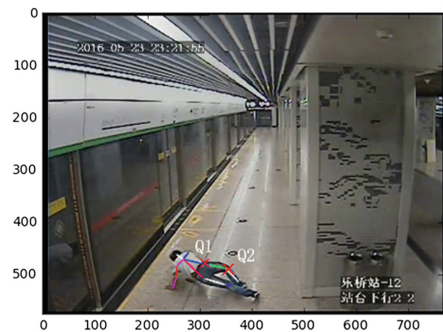


Fig. 2. Calculation points of leg angle

legs are calculated respectively. The formula for calculating the angle value of one leg is as follows:

$$\theta = \tan^{-1}[(y_{Q2} - y_{Q1}) \div (x_{Q2} - x_{Q1})] \tag{3.1}$$

$x_{Q1}, y_{Q1}, x_{Q2}, y_{Q2}$ were calculated as the horizontal and vertical coordinates of the key points of the hip and knee on one side of the body. The specific instructions are shown in Fig. 2.

Table 1. Angle statistics of six figures

left and right leg angle	people							
Figures	1	2	3	4	5	6	7	8
Fig.3	[93.37,91.74]	[-32.91,83.66]	[92.29,100.49]	[158.15,156.47]	[75.96,0.0]	[107.93,90.0]	158.58,92.49]	[87.61,354.29]
Fig.4	[91.51,93.27]	[98.53,101.89]	[64.23,85.76]	[95.19,95.19]	[70.71,59.93]	null	null	null
Fig.5	[88.49,93.27]	[97.60,99.78]	[98.13,92.73]	[90.0,103.67]	[87.14,98.53]	[95.19,100.78]	[85.24,-11.56]	[95.71,90.0]
Fig.6	[90.0,91.64]	[86.99,0.0]	[99.46,105.95]	[90.0,86.99]	[45.81,109.98]	[90.0,96.12]	null	null
Fig.7	[96.34,95.19]	[94.24,90.0]	[90.0,83.66]	[88.41,14.28]	[90.0,100.78]	[95.19,120.65]	null	null
Fig.8	[97.50,91.74]	[85.43,92.20]	[27.82,25.71]	[107.35,112.75]	[92.73,100.78]	null	null	null

Six different pictures, each containing a fallen person, were skeletally extracted to calculate the angle of the person’s left and right thighs. The green and white lines represented the person’s left and right thighs respectively. As shown in Table 1, the left and right leg angles with corresponding serial numbers in each drawing are listed (the persons in each drawing are independently numbered), corresponding to Figs. 3, 4, 5, 6, 7 and 8 respectively, in which the red data is the angle values of the fallen persons’ two legs.



Fig. 3.



Fig. 4.



Fig. 5.



Fig. 6.



Fig. 7.



Fig. 8.

It is found that the leg angle of normal erect pedestrians is about 90° and that of fallen pedestrians is quite different from 90° .

Result Evaluation. Using the abnormal angle of the leg for judging falling accords with the general law. It can accurately determine whether the person in the picture is in the abnormal attitude of falling or normal standing posture.

3.2 Falling Track Based on Continuous Frames

The original human skeleton extraction algorithm PAFs (part affinity fields) [4] uses frame-by-frame real-time processing for video without human tracking. Considering that the fall behavior includes three states: before fall, during fall, after fall, the corresponding state of leg angle will be from normal to abnormal, finally to normal, if such information can be obtained from the video frame, it can be more accurate to determine the presence of a fall in this surveillance video.



Fig. 9. Tracking through key coordinates of the neck in consecutive frames

As shown in Fig. 9, in the four consecutive frames, the necks of A and B are marked in red. It can be seen that the distance between the adjacent frames is very short. In the current video frame, if there is a neck coordinate in the neighborhood of the neck key point coordinate in the previous frame, the two will be judged to belong to the same person.

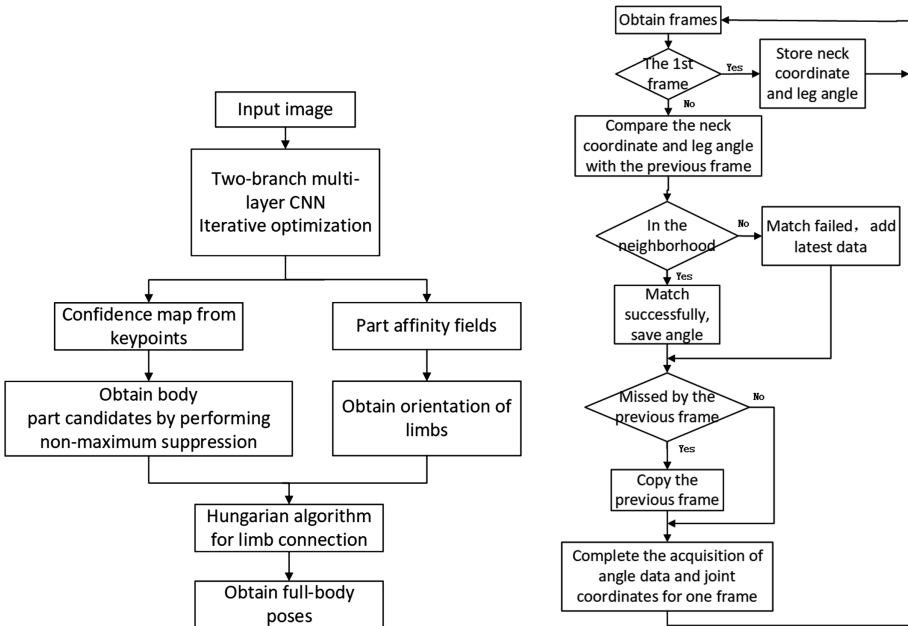


Fig. 10. Algorithm flow of PAFs [4]

Fig. 11. Algorithm flowchart

Based on PAFs (part affinity fields) [4], this paper extracts the key points of human body and the information of limb connection. The algorithm flow of PAFs (part affinity fields) [4] is shown in Fig. 10. After that, the human is tracked by the position coordinates of the neck, and the fall behavior is monitored from the tilt angle of the thigh.

Through cyclic reading video stream, the above method can track the same person, and store the angle of each frame of each person, then estimate the attitude by calculating the average angle of each frame in real time for a period of time. If numerical abnormalities occur repeatedly, it is judged that a fall occurs in the surveillance video. The algorithm flowchart is shown in Fig. 11.

4 Experimental Results

4.1 Single Frame Fall Judgment

Recognition Criteria. Before recognizing a fall, we should analyze the difference between the body's thigh angle value in normal and fall. Figure 12 is a few pictures containing these two postures, and Fig. 13 is the leg angle data of all the video frames corresponding to the above postures. It can be found in the picture that the person is in a fall from more than 50 frames to more than 300 frames. And the leg angle from the scatter plot in Fig. 13 can be found to deviate from 90 while the value of the other frames is around 90.



Fig. 12. Falling keyframe

By analyzing the characteristics of angle values in different states, this paper determines the classification rule of behavior in a single picture as follows:

(a) normal

$$|\text{leg angle} - 90^\circ| < 30^\circ;$$

(b) fall down

$$|\text{leg angle} - 90^\circ| \geq 30^\circ.$$

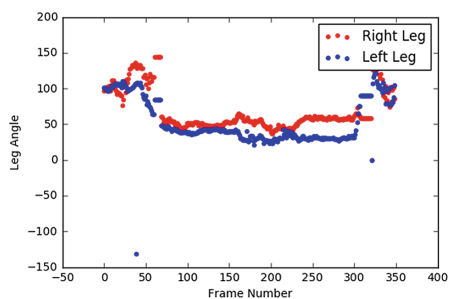


Fig. 13. Leg angle of all frames

Results. In this paper, 76 pictures were selected from fall videos, including 60 fall pictures and 16 normal posture pictures. Some pictures contain a single person and some have a number of people. On the premise of extracting human skeleton correctly, there are 60 fallen people and 117 normal people in these 76 pictures. Based on the above criteria, the final results are summarized as shown in Table 2.

Table 2. Gesture recognition results

Actual posture	Decision posture		
	Judged as fall	Judged to be normal	Accuracy
Falling	50	10	Fall detection 83.33%
Normal	7	110	Error alarm 5.98%

Result Analysis

(1) fall is judged to be normal

The main reasons for falling judged as normal posture are:

(a) the particularity of falling posture.

As shown in Fig. 14, the falling person’s upper body bends, while the lower limbs are still upright. This paper judges the fall on the basis of legs bending, so it can be misjudged.

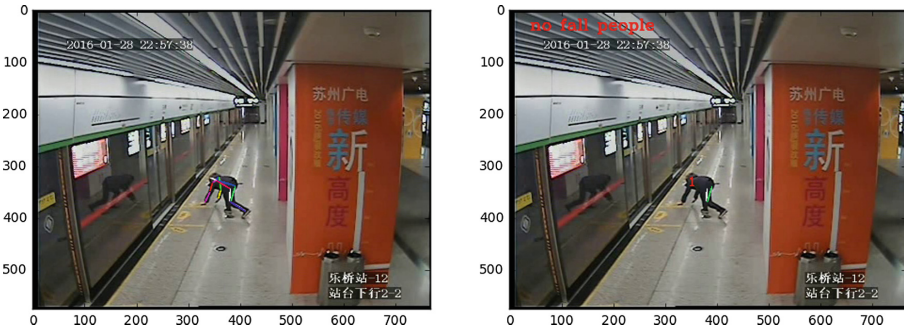


Fig. 14. Special fall posture

(b) the particularity of fall direction.

As shown in Fig. 15, the human body is in the same direction as the surveillance shooting, resulting in this method cannot distinguish the change of leg angle in this direction, although his two legs have changed from vertical to horizontal. So, it cannot identify the fall.

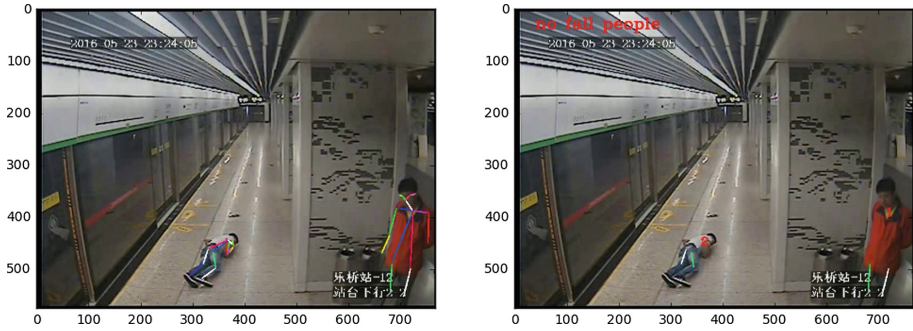


Fig. 15. Special fall direction

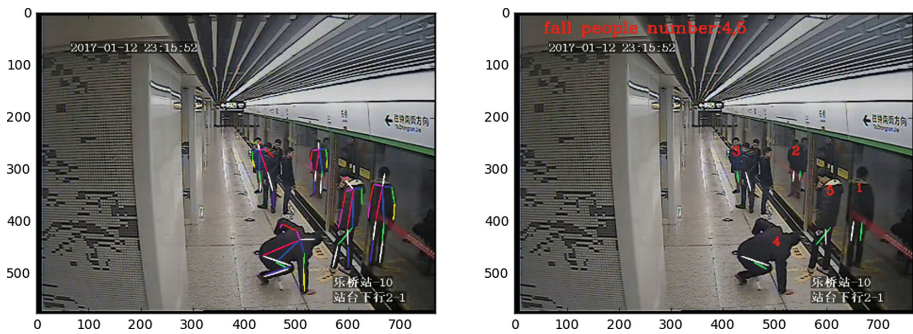


Fig. 16. Normal people misjudged

(2) normal posture is judged as fall

The main reasons for normal posture misjudged are:

(a) the particularity of normal standing posture

As shown in Fig. 16, the leg of standing person (marked as 5) is not completely upright, resulting in misjudgment. However, this attitude is very common in daily life, so error rate of relying only on a single picture for fall recognition is high. In contrast, determining from multiple pictures and multi-frame data is more reliable. Detailed results about real-time fall determination can be seen in the next section.

(b) imperfect posture recognition model

The attitude recognition algorithm is not perfect enough. As shown in Fig. 17, due to the depth of platform, the judgment of the leg angle of the person standing in the distance is wrong, which results in the algorithm judging the person wrongly as a fall.

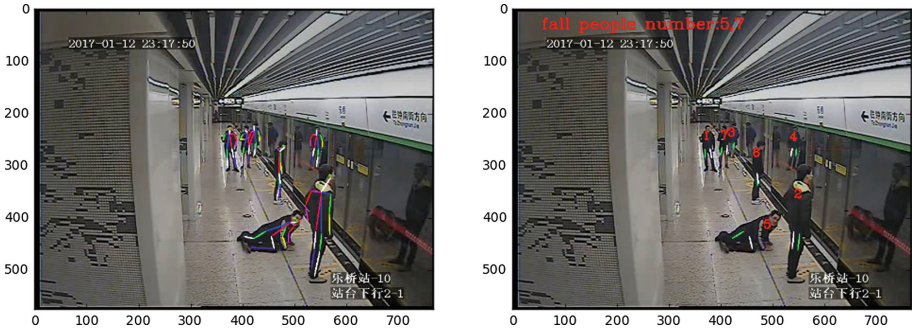


Fig. 17. Distant people misjudged

4.2 Real-Time Fall Judgment Through Continuous Frame

Judging a fall based on a single picture lacks fault tolerance, and some occasional special gestures can easily affect the final result. Therefore, this section analyzes the video frame by frame, and estimates the attitude by tracking the neck key points. In this paper, two videos were tested. The results show that the alarm can be sent out within two seconds of a fall, and the fall event can be accurately determined in the current video.

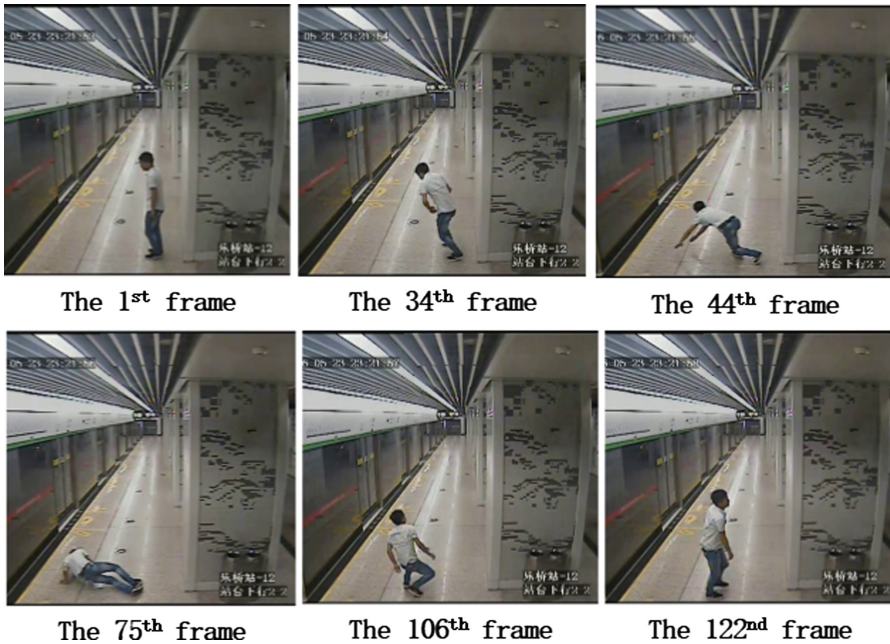


Fig. 18. Falling keyframes in the 1st video

Figure 18 is several key frames in the first video that containing falling. The corresponding leg angle values are shown in Fig. 19. It can be seen that the whole fall process is from about 40th frames to 150th frames.

In this paper, the human leg data is judged in real time, and the alarm is sent in two seconds. As shown in Fig. 20(a), real-time decision data shows that an early warning is issued in frame 100, that is, a decision is made successfully by relying on the first 100 frames of data.

Figure 20(b) is the 100th frame, and the person labeled 1 is decided as falling.

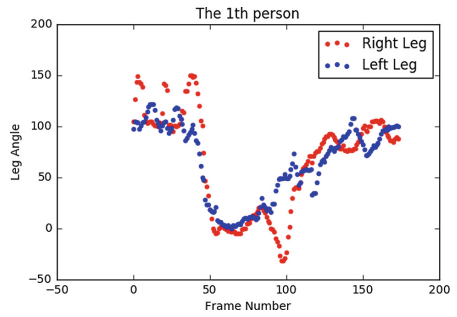
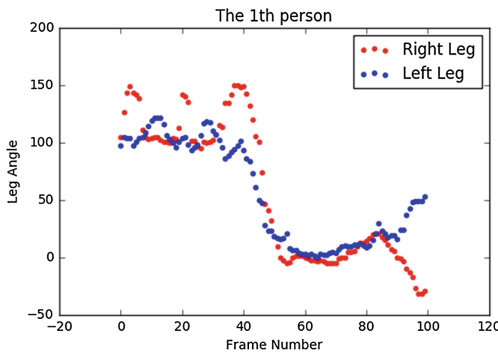


Fig. 19. Complete date

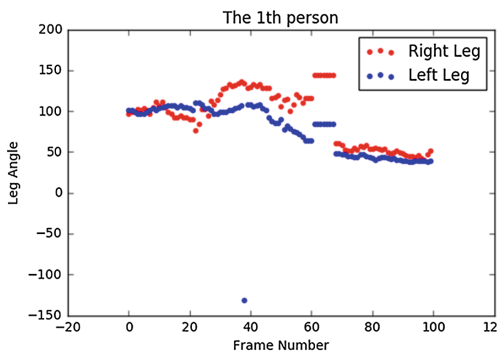


(a)



(b)

Fig. 20. Real-time decision data



(a)



(b)

Fig. 21. Real-time decision data

The key frames in the second fall video are shown in Fig. 12. Figure 13 is the angle data read frame by frame from the whole video. As can be seen from the scatter plot, the people in the video is in the fall posture from about the 50th frame to the 300th frame. Figure 21(a) is the data obtained before a fall event is determined. It can be seen that the abnormal fall behavior in the video can be identified based on the first 100 frames. Figure 21(b) is the video frame when the alarm is issued at the corresponding time. The figure shows that the person labeled 1 has fallen.

5 Summary

In this paper, the threshold method is used to locate the neck key point of the same person, so as to achieve the tracking of the characters in the continuous video frames. Making full use of each frame of the video information, the real-time analysis can alarm after two seconds of fall. The bending of the human thigh, that is, the angle between the leg and the horizontal plane, is taken as the feature for the fall judgment, and the threshold value is determined by experience. Simplicity and certain accuracy are guaranteed at the same time. However, due to the limitation of the original attitude recognition algorithm in resolving the occlusion of personnel, the occasional occurrence of falls, and the limitation of the characteristics of falls judgment, there are inevitable errors in the process of image processing, resulting in the inconsistency of the number of people in the front and rear frames, the errors of key point coordinates and wrong limb connection. From the research point of view, the fall data used for training is insufficient. Different scenes, different shooting angles, static and dynamic camera data are lacking, so it is not adequate enough to verify this algorithm's accuracy. There remains many improvements in testing effects and optimizing the recognition algorithm.

References

1. Tompson, J.J., Jain, A., LeCun, Y., et al.: Joint training of a convolutional network and a graphical model for human pose estimation. In: *Advances in Neural Information Processing Systems*, pp. 1799–1807 (2014)
2. Aslan, M., Sengur, A., Xiao, Y., et al.: Shape feature encoding via fisher vector for efficient fall detection in depth-videos. *Appl. Soft Comput.* **37**, 1023–1028 (2015)
3. Ke, S.R., Thuc, H.L.U., Lee, Y.J., et al.: A review on video-based human activity recognition. *Computers* **2**(2), 88–131 (2013)
4. Cao, Z., Simon, T., Wei, S.E., et al.: Realtime multi-person 2D pose estimation using part affinity fields. In: *CVPR 2017*, vol. 1, no. 2, p. 7 (2017)
5. Foroughi, H., Pourreza, H.R.: Intelligent video surveillance for monitoring fall detection of elderly in home environments. In: *11th International Conference on Computer and Information Technology* (2008)
6. Rougier, C., Meunier, J., St-Arnaud, A., et al.: Fall detection from human shape and motion history using video surveillance. In: *2007 21st International Conference on Advanced Information Networking and Applications Workshops, AINAW 2007*, vol. 2, pp. 875–880. IEEE (2007)



Determining Acceleration Lane Length on Expressway Weaving Area Using Microscopic Traffic Simulation

Jing Luo^{1,2}, Xiaoqing Zeng³, Yinong Chen⁴,
and Daniel(Jian) Sun^{1,2(✉)}

¹ State Key Laboratory of Ocean Engineering,
School of Naval Architecture, Ocean and Civil Engineering,
Shanghai Jiao Tong University, Shanghai 200240, China
carryonex@sjtu.edu.cn

² Center of Intelligent Transportation System, UAV Applications,
Shanghai Jiao Tong University, Shanghai 200240, China

³ Traffic Information and Control Joint Center, School of Transportation
Engineering, Tongji University, Shanghai 201804, China

⁴ Computer Science and Engineering, School of Computing,
Informatics, and Decision Systems Engineering (SCIDSE),
Arizona State University, Tempe, AZ 85287-8809, USA

Abstract. As the inter-junction section of the mainline and the ramps, the expressway weaving area is one of the major bottlenecks within the urban expressway system. In addition to various behaviors of drivers, the acceleration lane length of expressway weaving area is generally short, which makes it more prone to trig traffic disorder and become a traffic accident blackspot. Increasing the acceleration lane length of expressway weaving area can provide larger buffer space for the merging process of vehicles. However, a longer length generally means a higher infrastructure investment. Consequently, an optimal length of acceleration lane is essential for the expressway road infrastructure development. This paper uses the empirical survey data and a microscopic simulation system, called Traffic Parallel Simulation System (TPSS), as the simulation platform to build and verify the lane-changing model in expressway weaving area. By means of the developed weaving area simulation model, the impacts of various acceleration lane length on the traffic flow service levels are investigated, for obtaining the recommended value of acceleration lane length under different levels of services. This value further provides the guidelines for traffic management departments to formulate strategies and technical supports to relieve traffic congestions.

Keywords: Expressway weaving area · Acceleration lane length · Microscopic traffic simulation · Design method

Foundation item: Project (16ZDA048) supported by the National Social Science Foundation of China.

1 Introduction

As a typical facility of urban road system, expressways play an important role within traffic corridors of medium and large capacities. Due to the comparable high traffic demand, the level of services of expressway at peak hours remains inferior in many large cities (Research institute of highway ministry of transport 2000). As the interjunction of the mainline and the ramp, the expressway weaving area is generally with a large amount of lane-changing maneuvers, and the frequent lane changing may lead to traffic disorder and deteriorated road capacity. It is also prone to trigger serious traffic congestion, especially when the mainline traffic is close to the saturation state, in which the traffic capacity of merging area is reduced, making it the major bottleneck of expressway. Consequently, it is urgent and essential to reduce the influence of the bottleneck in merging area and to improve the traffic capacity, and thus to maintain service of the entire expressway system.

Expressway merging area generally consists of three components: the ramp, mainline, the acceleration lane, and the mainline (Ministry of Transport of the People's Republic of China 2006). The acceleration lane directly determines the duration for vehicles to complete the lane-changing maneuver. It is an important parameter affecting the combined flow operation characteristics. At present, studies on the length design for the acceleration lane are based on literatures (Shao 2001) formulating the aspect of statistical probability of vehicles entering into highway on the acceleration lane and obtaining a practical model to determine the length of the acceleration lane. Zhao (2004) formulated the ramp capacity of the main line under different traffic volumes and established a probability model in which the on-ramp vehicles could directly merge under different traffic loads on the main road. The study also provided the setting rules of the acceleration lane length under certain ramp traffic probability and the traffic volume of the main road served. Zhi (2009) formulated the probability of vehicle convergence on the acceleration lane. The designed expressway lane length was optimized based on the traffic flow of the main highway, the acceleration lane, and the minimum time headway. In this way, traffic flow control model of the entrance ramp was established. Yang (2015) formulated a model that include most of the existing studies on acceleration characteristics. However, this model does not fully represent the realistic traffic conditions or are based on outdated data. This study investigated the actual acceleration characteristics at metered on-ramps according to field data collection and analysis. Accordingly, the distance-acceleration profiles of various ramp configurations were determined. Yang (2016) developed a method for determining acceleration lengths at the metered on-ramps. They built a range-velocity profile for each ramp and developed regression models to predict the required acceleration length at a given merging speed.

In recent decades, with the rapid development of computer simulation and intelligent transportation technologies, traffic simulation was introduced to analyze the operating characteristics of expressway weaving area, which has facilitated the emerge of various simulation packages. A number of microscopic traffic simulation models have been developed, including VISSIM (Schneeberger 2003), AIMSUN (Barceló et al. 1998), PARAMICS (Lu et al. 2010), MITSIM (Yang and Koutsopoulos 1996), CORSIM (ITT Industries 1999) and SITRAS (Hidas 2002), and so on. However, as the

construction basis of such simulation models are in accordance with the unique vehicle, roads, drivers and other traffic flow features in different countries, it does not necessarily match China's road traffic characteristics (Sun et al. 2013). In addition, the factors of the models, such as closure, specificity, and exclusiveness, result in inadequate application scope of such models. In the studies, there has been little simulation research on traffic operation in weaving area in China. Therefore, this paper uses a microscopic traffic simulation package TPSS, jointly developed by Shanghai Jiao Tong University and Jilin University as the simulation platform, to implement lane-changing model in weaving area for the development of a simulation system. The system is developed using Visual C++ programming language.

The simulation model of weaving area was introduced to analyze the influence of different acceleration lane lengths on the traffic flow service level (Xiong and Wang 2016), thus to obtain the recommended value of acceleration lane length under different service levels. RMSE and MPE evaluation indices were selected to verify the validity of the model. Results of this study may provide the basis for traffic management departments to formulate further strategies and technical supports to ease traffic congestion.

The rest of the paper is organized in the following sections. Section 2 presents the major factors used in the simulation model construction in weaving area. Section 3 formulates the acceleration length model and establishes simulation experiment. Section 4 analyzes and discusses the simulation results and recommends the speed reference value of acceleration lane. Conclusions and recommendations for future work are provided in Sect. 5.

2 Simulation Model Construction

Traffic Parallel Simulation System (TPSS) is a time step based microscopic traffic simulation package (Juan 2008), simulating the dynamic characteristics of the large-scale network traffic flow. The system can be divided into three layers. The first layer is a graphic control workstation, namely, a foreground layer, composed of a large number of graphic workstations, providing a friendly and easily operating user interface. Windows operating system and Visual C++ are used to develop this layer. The main responsibilities of this layer include interface optimization and high maneuverability. This layer is not responsible for the stability and safety in the system design. The second layer consists of three components: a database server, a data center server, and a parallel computing dispatch server. The main function is to provide communication and scheduling for the frontend graph control station and the background parallel computing workstation. This layer is mainly responsible for the safety and stability of the system, and Linux is adopted instead of Windows. The third layer consists of multiple parallel computing workstations and is the background layer. The main function is to receive the commands from the second layer and execute the corresponding tasks, with the computing results of parallel workstations sending back to the second layer. This layer is also deployed on a highly safe and stable Linux operating system.

2.1 Modeling Weaving Area

In the traffic simulation network that we design, five entities are designed and included in the expressway weaving area: road section, lane, ramp, vehicle and driver. According to the object-oriented design principle (Eldredge 1990), these entities are encapsulated as classes as TLink, TLane, TRamp, TVehicle, and TDriver. Combined with the related entity attributes and call requirement, these classes that characterize expressway weaving area simulation are defined for member functions and property variables. Five vector queue structures were defined accordingly, as below.

- (1) TVEHICLELIST: represents the vehicle vector queue structure, used to store the class objects of all vehicle generated in the simulation network, such as the types of vehicle, displacement, speed, acceleration, as well as the front and rear clearance of lane changing, and the number of lane-changing, etc.
- (2) TLINKLIST: represents the section vector queue structure, used to store the class objects of all sections during the simulation network, including the basic properties of sections such as serial number, number of lanes, number of lane channelization, grade, and slope as well as the road running conditions such as flow, density, speed, delay time, etc.
- (3) TLANELIST: represents the lane vector queue structure, used to store the class objects of all lanes in the section, including the lane basic properties such as width, length, and control mode, as well as the lane operation conditions such as flow, speed, density, queue length, and number of stops, etc.
- (4) TRAMPLIST: represents the ramp vector queue structure, used to store the class objects of all ramps, including ramp type, acceleration lane length, gradient length and so on.
- (5) TDRIVERLIST: represents the driver vector queue structure is used for the class objects of driver's characteristics, including the driver's type, the normal distribution of the mean and variance, the driver's reaction time, and the acceptable factor for the speed limit of the road, etc.

Based on the definition of these data structures, the simulation modules can share the dynamical information without affecting each other, so that the reading efficiency and speed are significantly increased during the simulation, which is especially important for large-scale network simulations.

2.2 Lane-Changing Module Design

According to the division of lane-changing maneuvers and the interactions between the mainline vehicles and the merged vehicles in expressway weaving area, the lane-changing modules are redesigned and divided into five sub modules, as follows:

Free lane change: It mimics the merged vehicles to adopt free lane changing, namely, the lane-changing behavior under light traffic flow in expressway weaving area. By computing the gap acceptance in free change situation, whether the merged vehicles can adopt free type lane change are evaluated. If the condition is satisfied, the free lane change behavior is executed. Otherwise, it is necessary to continue to judge

whether the subsequent collaborative lane change and forced lane change can be performed.

Cooperative lane change: It mimics the cooperative behavior by the lag vehicles in the mainline, specifically including lane change collaboration and deceleration collaboration. This module first evaluates the probability of following vehicle to adopt change lane collaboration and deceleration collaboration. If the condition is satisfied, the collaborative lane change is performed; otherwise, the merging vehicle has to judge whether it is necessary to adopt forced lane change.

Force lane change: It mimics the forced lane change behavior by the merged vehicles. If a merged vehicle finds that the conditions for free lane change and collaborative lane change are not satisfied, it will consider whether to force the following vehicle to decelerate in order to merge into the mainline. This module evaluates the probability to adopt force lane change. If the condition is satisfied, the force lane change will be performed, otherwise, the merging vehicle has to accelerate or stop.

Gap acceptance: It determines the feasibility for vehicle to change lanes. Before performing the lane changes, this module should be invoked to decide the probability to change lanes, including the judgement of three lane changes maneuvers, i.e. free lane change, collaborative lane change and forced lane change.

Merging-in: If the gap acceptance model satisfies the corresponding requirement, the vehicle performs the merging in maneuver.

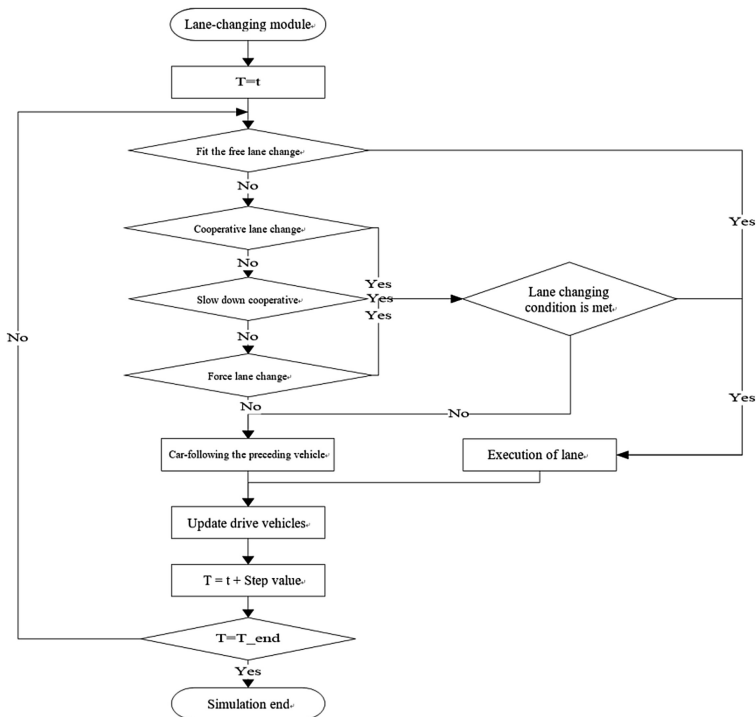


Fig. 1. Lane changing for the merging maneuver on expressway weaving areas.

As a merging vehicle must not interfere with the mainline vehicles during the free lane change process, it first judges the free lane change after entering acceleration lane. If the condition is satisfied, the merged vehicle will perform a free lane change, and the original guiding vehicle on the mainline will be set as the front vehicle. Otherwise, the collaborative merging or forced merging will be considered. The flow chart for the lane changing module is provided (see Fig. 1).

3 Modeling Length of Acceleration Lane

3.1 Determining Service Volume

In the expressway merging area, the traffic volumes on the mainline and on ramp not only play a role in determining the characteristics of traffic flow, but also the fundamental parameters for microscopic traffic simulation. In accordance with Design Specifications of Urban Expressway (CJJ (37-2012)) (Code of design for urban and engineering 2012), the service level of expressway is generally divided into four grades, as presented in Table 1. As the urban roads are in a dense network system and the basic section is relatively shorter than the regular highway, it is necessary to apply three-grade level of services to set the design speeds of the mainline of expressway at 60 km/h, 80 km/h, 100 km/h, respectively, as in Table 2. It is easy to determine the corresponding service levels of basic sections according to different speeds. The maximum service traffic volume of the upper grade is the minimum service traffic volume of the next grade. Accordingly, the service traffic volume ranges of different speeds under the three-grade level are determined.

Table 1. Level of service classification for expressway under 100 km/h.

Level of service classification	Traffic density [pcu/(km·lane)]	Speed (km/h)	V/C	Maximum service volume [pcu/(h·lane)]
Level 1 Free flow	≤ 10	≥ 88	0.40	850
Level 2 Steady flow	≤ 20	≥ 76	0.69	1500
Level 3 Steady flow	≤ 32	≥ 62	0.91	2000
Level 4 Saturation flow	≤ 42	≥ 53	Close to 1.00	2200

Table 2. Designed traffic volume under three levels of mainline traffic.

Design speed (km/h)	Mainline service volume of traffic (lower limit) [pcu/(h·lane)]	Mainline service volume of traffic (upper limit) [pcu/(h·lane)]
100	1500	2000
80	1300	1600
60	1000	1400

The identified service traffic is mainly for the ideal situation, which is necessary to be amended based on related influencing factors in field traffic situations to obtain the feasible values. Considering the influence of medium and large vehicles, the reduction coefficient of large vehicles is approximated as (Ministry of Transport of the People's Republic of China 2006):

$$f_{HV} = \frac{1}{1 + \sum p_i(E_i - 1)} \quad (1)$$

where,

f_{HV} — the reduction coefficient of large vehicle;

p_i — the percentage of traffic volume of i type vehicle from the total;

E_i — the reduction coefficient of i type vehicle.

The conversion coefficient is chosen as 1, for light-duty and mid-size vehicles, and 1.5 and 2 for large vehicles and articulated buses, respectively.

In accordance with a large amount of field investigation of expressway merge areas in Shanghai, China, small-type vehicles are dominant in the expressway merge areas, which accounts for 93.1% of the total traffic volume, while large-type and medium-type vehicle account for 2.3% and 4.6%, respectively. According to the conversion coefficient of different vehicle types stipulated by CJJ (37-2012), the conversion coefficient of large-type vehicles is set as 0.97, based on which the mixed traffic volume of different designed speeds can be calculated. A six-lane expressway was used as an example for the service traffic volume calculation, as most expressways in China have six lanes in both directions. On the design speed 100 km/h, the single lane volume of traffic is 1500–2000 pcu/(h·lane), and the correction factor is 0.97. Therefore, the mixed traffic volume can be approximated as 1455–1940 vehicle/(h·lane), and the overall three-lane volume is set at 4365–5820 veh/hr.

To achieve the three-grade service of the main lane of expressway, it is necessary for the traffic volume of ramp to realize three-grade service. According to CJJ (37-2012), the speed of ramp has to be designed between 30–40 km/h. The traffic capacity of ramps with different designed speeds is 1050 pcu/(h·lane) under 20 km/h, 1133pcu/(h·lane) under 30 km/h, 1196 pcu/(h·lane) under 40 km/h.

It can be determined that when the designed speed reaches 30–40 km/h, the basic traffic capacity of ramp is 1133–1196 pcu/(h·lane). As the value range of density is 0.5–0.8 below three-grade service, it was calculated that the service traffic volume of ramp is 587–956 pcu/(h·lane) for the designed speed of 40 km/h, and the service traffic volume of ramp is 573–920 pcu/(h·lane) for the designed speed of 30 km/h.

When the ramp traffic merged in, the vehicle spacing is influenced by the main lane traffic volume, so that the permissive the ramp traffic volume merging-in can be approximated based on the main lane traffic volume. In the expressway merge area, the characteristics of ramp traffic changing lanes to merge into the main lane is similar to that secondary road traffic running at no signal control intersections. Therefore, the computing method of secondary road traffic volume at no signal control intersections is used. Within the traditional distribution models describing the time headway, Erlang distribution was used to model the main lane time headway as follows.

The probability density function of Erlang distribution is:

$$p(t) = \frac{\lambda(\lambda t)^{r-1}}{(r-1)!} e^{-\lambda t}, r = 1, 2, \dots, n \quad (2)$$

The distribution function of Erlang distribution is shown as the following:

$$p(h \geq t) = \sum_{i=0}^{r-1} \frac{(\lambda t)^i}{i!} e^{-\lambda t} \quad (3)$$

The mean value and variance of the distribution are as follows:

$$M = \frac{r}{\lambda}, D = \frac{r}{\lambda^2} \quad (4)$$

where r is the order of Erlang distribution; when $r = 1$ or $r = \infty$, Erlang distribution can be simplified as the negative exponential distribution or to produce mean time headway. It can be found that the different values of parameter r can cause different time headway distribution under different running state, which can successfully show different traffic flow characteristics (Chang et al 1998).

When the time headway of main lane traffic follows the Erlang distribution of r order, the ramp traffic capacity is estimated as follows:

$$Q_r = Q_1 \sum_{i=0}^{r-1} \frac{(-1)^i}{i!} \left| \frac{\exp(-Q_1 r t_0 x / 3600)}{1 - \exp(-Q_1 r t x / 3600)} \right|_{x=1}^{(i)} \quad (5)$$

where:

- Q_r — ramp traffic volume;
- Q_1 — main lane traffic volume;
- t_0 — merging probability;
- t — critical gap of ramp vehicle;
- r — the order of Erlang distribution.

The ramp service traffic volume can be calculated from Eq. (7). For the mainline design speed as 100 km/h, the field service volume of traffic is approximated as 551–777 veh/hr; for the mainline design speed as 80 km/h, the field service volume of traffic is approximated as 723–880 veh/hr; for the mainline design speed as 60 km/h, the field service volume of traffic is approximated as 826–907 veh/hr.

3.2 Simulation Case Design

The basic parameters of traffic flow, road geometry and vehicle characteristics are explained as below.

3.2.1 Traffic Flow Parameters and Road Geometry

The traffic flow parameters mainly include the parameters for the mainline and ramp traffic under different designed speeds, vehicle type composition, and so on, while the

road geometrical parameters include number of lanes, width of lane, length of lane and lane type. The mainline and ramp traffic for simulation model is input according to the values identified in following Sect. 4.1. The vehicle type composition is according to the investigation of nine merge are as in Shanghai, in which small-type vehicle accounts for 93.1%, while the medium and large-type vehicle account for 4.6% and 2.3%, respectively. The length of acceleration lane is set between 60–260 m to evaluate the influence of different lengths of acceleration lane on the corresponding service level (Table 3).

Table 3. Input parameters of traffic flow and road for TPSS.

Parameter	Input value	
Mainline traffic volume (veh/h.)	5500, 4600 and 4000	
Ramp traffic volume (veh/h.)	612, 760 and 868	
Traffic composition	Light-duty vehicle	93.1%
	Mid-size	4.6%
	Large vehicle	2.3%
Number of main lanes	3	
Number of ramp lanes and acceleration lanes	1	
Length of acceleration lane (m)	60, 80, 100, 120, 140, 160, 180, 200, 220, 240, 260	

3.2.2 Vehicle Characteristic Parameters

The parameters for the vehicle characteristics mainly include the expected speed distribution, the acceleration and deceleration. The expected speed is important to traffic capacity and running speed. In order to reasonably calculate the traffic performance of merge area under different designed speeds and levels of service, the expected cruising speed is defined in accordance with the required values under three-grade service specified in CJJ (37-2012). Thus, the minimum and maximum expected speeds are determined. Parameters of expected speed distribution under different designed speeds are 62–76 km/h under 100 km/h design speed; 54.5–64 km/h under 80 km/h design speed; 43.5–50 km/h under 60 km/h design speed.

For further evaluation and analysis, simulation result including several indicators of traffic, speeds of different sections of merge area, and density were output. From randomness purpose, each simulation case was run three times, with the average values used.

3.3 Impact of Accelerating Lane Length

Results from different simulation cases are obtained to assess the influence of acceleration lane length on the level of service of the merge area. Considering a main lane with designed speed of 100 km/h and the ramp with designed speed of 40 km/h, the influence on the upstream of merge area, as well as the merge area were investigated.

The relation between length of acceleration lane and running speed of upstream is shown (see Fig. 2). It can be seen from that the running speed of upstream of merge area has positive correlation with the length of acceleration lane. When the length of acceleration lane increases to about 205 m, the service level of upstream of merge area increases from four-grade to three-grade (the running speed is more than 62 km/h); When the length of acceleration lane increases to larger than 210 m, the speed of upstream of merge area tends to be stable.

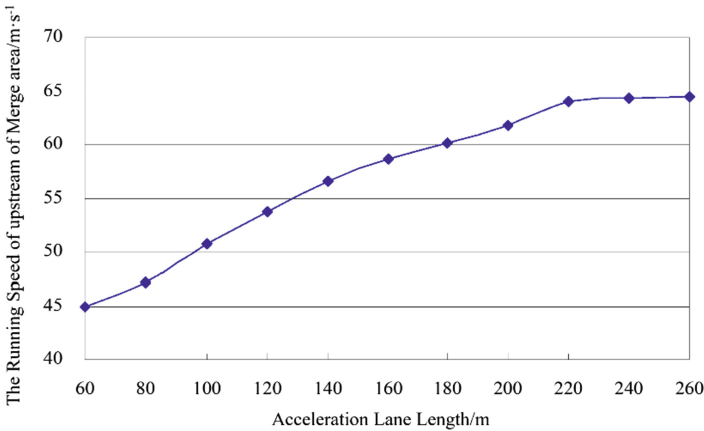


Fig. 2. The running speed of upstream of merge area.

The relation between the length of acceleration lane and running speed of merge area is shown (see Fig. 3). The running speed of merge area has positive correlation with the length of acceleration lane. When the length of acceleration lane increases to about 210 m, the service level of merge area increases from four-grade (forced flow) to four-grade (saturated flow) and the running speed reaches 53 km/h. When the length of acceleration lane increases to more than 220 m, the length of acceleration lane has no effect on the increase of speed and the curve of running speed of merge area tends to a horizontal line.

Based on the analysis of influence of the length of acceleration lane on the upstream of merge area, as well as merge area, to guarantee the three-grade service of main lane and four-grade (saturated flow) services of the merge area, the length of acceleration lane is supposed to be no less than 220 m.

Analysis was carried out on the influence of length of acceleration lane under speed combinations of 100–30, 80–40, 80–30, 60–40, and 60–30 with the same analysis approach to identify the required minimum length of acceleration lane based on the three-grade services of the mainline and the four-grade (saturated flow) services of the merge area.

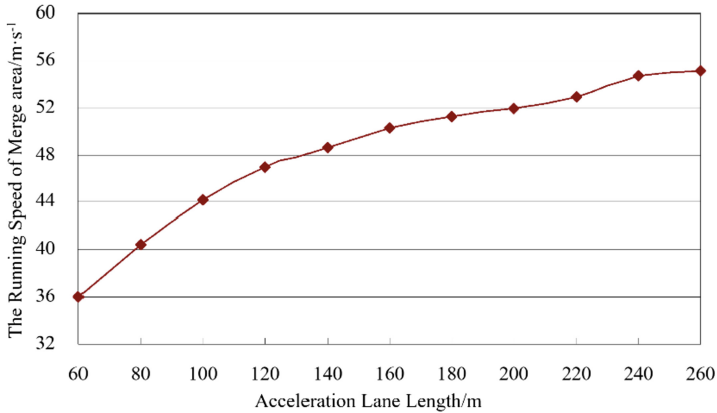


Fig. 3. The running speed of merge area.

4 Result Analysis and Discussion

4.1 Model Simulation and Validation

The model is implemented and validated using the expressway weaving area of Xinzhuang expressway interchange of Shanghai (see Fig. 4). The mainline traffic flow, acceleration lane traffic flow, the proportion of vehicles, the running speed and other related data were obtained based on the video processing, as shown in Table 4. To overcome the randomness of simulation, multiple runs of simulation were carried out, with the average value of outputs used.



Fig. 4. Simulation case output animation.

Table 4. The survey location data.

Indicators	Traffic composition	Mainline	Acceleration lane
Traffic composition (%)	Light-duty vehicle	91.8	92.8
	Mid-size	3.9	7.0
	Large vehicle	4.3	0.2
Speed (m/s)	Average	13.1	12.2
	Standard deviation	4.3	3.3
Traffic volume (pcu/h)		2027	536
Number of lanes		3	1

To validate the merging area simulation model, the lane change location of the merged vehicle, the lane change gap of following vehicle and other micro index were selected as the indicators, by referring to HCM2010 (Ryus et al 2010). Two evaluation indices, RMSE (Charisma 2002) and MPE (Pindyck and Rubinfeld 1997), were selected for comparing and analyzing simulation and observed values, as below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (M_i^{sim} - M_i^{obs})^2} \quad (6)$$

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{M_i^{sim} - M_i^{obs}}{M_i^{obs}} \quad (7)$$

where:

n — survey data sample;

M_i^{sim} — simulation value of merging location;

M_i^{obs} — measured value of merging location;

$i = 1, 2, 3 \dots, n$

Through the observation of simulation animation, the simulation model of the merge area of the expressway can well represent the complex interaction behavior between the mainline vehicles and the incoming vehicles in the merge area. Some mainline vehicles will take the way of deceleration or lane change to cooperate with the incoming vehicles to lane change and merge into the mainline. Then, the average deviation and the average percentage deviation of the two evaluations are evaluated against the simulation output and the measured data against the actual data, and the error of the estimate. The RMSE and MPE of the two indexes, e.g. the merged location, the gap of following vehicle, were obtained. The comparison and analysis of simulation and measured values were provided (see Fig. 5).

The result demonstrated that the average percentage deviation of the lane-changing location, the gap of following vehicle are 7.5% and 7% respectively, both within an acceptable range of 10%. In addition, the root-mean-square deviation of lane change location of merged vehicle, the lane change gap of the following vehicle was also smaller. Consequently, it can be figured out that the simulation output and the measured value have high consistency, indicating validity of the lane-changing model of

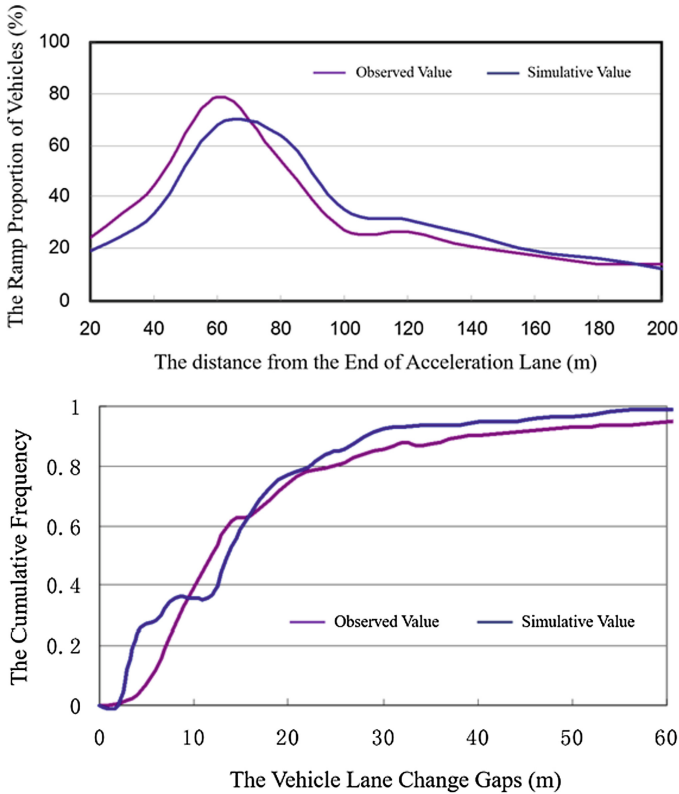


Fig. 5. Comparison of measured and simulation values of the lane-changing position and gaps.

expressway weaving area, which can better replicate the complex operation of the traffic flow in merging area.

4.2 Determining Length of Acceleration Lane

Based on the combination of traffic level of service and the designed speeds of mainline and ramp, the influence of the length of acceleration lane on the running state of the upstream of merge point and the merge area were assessed. The target is to keep the mainline service stays above three-grade and the merge area service stays above four-grade (saturated flow). The required minimum length of acceleration lane (single lane) is recommended as in Table 5 below.

From the comparison of the recommended minimum length of acceleration lane obtained in the simulation and the minimum value specified in CJJ (37-2012), it is found that the most designed length of acceleration lane is relatively short, as only the level of service of main lane is considered. Therefore, it is suggested that the length of the merge area should be extended in order to satisfy the requirements of levels of service for the upstream of the merge area, as well as the on-ramp merging area.

Table 5. Recommended minimum length of acceleration lane.

Ramp design speed (km/h)	Main lane design speed (km/h)		
	100	80	60
40	210 (180)	170 (160)	125 (120)
30	220 (180)	180 (160)	130 (120)

Note: 1. Values within the brackets is the value specified in Chinese Specification;
2. The length of acceleration lane excludes the length of transition section.

5 Conclusions and Recommendations

The length of acceleration lane in expressway merge area is essential in affecting the traffic flow characteristics in the merging area, and it plays a crucial role in determining the core content of expressway design. This paper proposed an expressway merging-in model based on the traffic volume of main line and ramp combined with the TPSS simulation system. We validated the simulation accuracy using the incorporated field data measured. RMSE and MPE indexes were calculated to indicate the reliability of simulation within certain acceptable range. Then, the reasonable values for expressway acceleration lane length were recommended, which may be observed directly from our animation when the acceleration lane length initially increases, which showed the speed of the merging area was enhanced clearly. When the length of acceleration lane attains a certain value, continuing to increase the acceleration lane length would reduce speed within the merging area. Therefore, simply increasing the acceleration lane length actually did not enhance the overall speed of the merging area, which may only significantly increase the construction costs. In future, we plan to combine the driver behavior characteristics with the facilities of expressway constructions, which can add additional optimization and ease traffic bottlenecks of urban expressway merging areas.

Acknowledgment. The authors are thankful to the Shanxi Bei'an Fire Protection Technology Co., Ltd. for providing the necessary data and information in this study. The research was sponsored in part by the Projects (17DZ1204003, 17ZR1445500) supported by the Shanghai Science and Technology Committee (STCSM), Shanghai Municipal Government, and the Projects (201705-JD-C1085-072, 201705-JD-C1085-056) supported by the National Independent Innovation Demonstration Zone. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

References

- Research institute of highway ministry of transport: Transport planning and research institute, and Southeast China University. Highway Capacity Research Report, China (2000)
Ministry of Transport of the People's Republic of China: Design Specification for Highway Alignment. China Communication Press, Beijing (2006)

- Shao, C.Q., Yang, Z.H., Chen, J.C., Ren, F.T.: A probability model for length of acceleration lanes on expressway. *J. Appl. Stat. Manag.* **20**(4), 42–45 (2001)
- Zhao, C., Deng, W., Zhou, R.G., Wang, W.: The length design method for acceleration lane taking into account of traffic. *J. Highw. Transp. Res. Dev.* **21**(8), 103–107 (2004)
- Zhi, Y.F., Zhang, J., Shi, Z.K.: Research on design of expressway acceleration lane length and merging model of vehicle. *China J. Highw. Transp.* **22**(2), 93–97 (2009)
- Yang, G., Xu, H., Tian, Z., Wang, Z., Zhao, Y.: Acceleration characteristics at metered on-ramps. *Transp. Res. Record J. Transp. Res. Board* **2484**, 1–9 (2015)
- Yang, G., Tian, Z., Xu, H., Wang, Z.: Recommendations for acceleration lane length for metered on-ramps. *Transp. Res. Record J. Transp. Res. Board* **2588**, 1–11 (2016)
- Schneeberger, J.D.: Microscopic simulation model calibration and validation: a case study of VISSIM for a Coordinated Actuated Signal System. In: *TRB 2003 Annual Meeting*, pp. 185–192 (2002)
- Barceló, J., Casas, J., Ferrer, J.L., et al.: Modelling advanced transport telematic applications with microscopic simulators: the case of AIMSUN Simulation technology. *Science and Art. In: Proceedings of the 10th European Simulation Symposium*, pp. 362–367 (1998)
- Lu, L., Yun, T., Li, L., et al.: A comparison of phase transitions produced by PARAMICS, TransModeler, and Vissim. *Intell. Transp. Syst. Mag.* **2**(3), 19–24 (2010). IEEE
- Yang, Q., Koutsopoulos, H.N.: A microscopic traffic simulator for evaluation of dynamic traffic management systems. *Transp. Res. Part C* **4**, 113–129 (1996)
- ITT Industries: *CORSIM user's manual*. Office of Safety and Traffic Operations Research and Development, pp. 1–37 (1999)
- Hidas, P.: Modelling lane changing and merging in microscopic traffic simulation. *Transp. Res. Part C Emerg. Technol.* **10**(5), 351–371 (2002)
- Sun, D.J., Zhang, L.H., Chen, F.: Comparative study on simulation performances of CORSIM and VISSIM for urban street network. *Simul. Model. Pract. Theory* **37**, 18–29 (2013)
- Xiong, Q.P., Wang, G.: An automatic electrical toll collection system for highway toll stations based on public transportation cards [China National Invention Patent], Patent ID: ZL2016.20131205.8, 28.9.2016
- Juan, Z.C., Tan, Y.L., Ni, A.N.: Study on microscopic traffic simulation model of public transport vehicles operation. *J. Highw. Transp. Res. Dev.* **25**(8), 119–127 (2008)
- Eldredge, D.L.: Applying the object-oriented paradigm to discrete event simulations using the C++ language. *Simulation* **54**, 83–91 (1990)
- Code of design for urban and engineering: Specification for design of urban expressway (CJJ 37-2012). China Architecture and Building Press, Beijing (2016)
- Chang, Y.L., Wang, W., Cao, H.: Capacity of inferior lane at unsignalized intersection. *J. Southeast Univ.* **28**(3), 98–102 (1998)
- Ryus, P., Mark, V., Lily, E., Richard, G.D., Barbara, K.O.: *Highway Capacity Manual (HCM2010)*. Transportation Research Board (2010)
- Charisma, F.C.: *Modeling Lane-changing Behavior in Presence of Exclusive Lanes*. Ph.D. Dissertation, Department of Civil and Environmental Engineering, MIT, U.S. (2002)
- Pindyck, R.S., Rubinfeld, D.L.: *Econometric Models and Economic Forecasts*, 4th edn. Irwin McGraw-Hill, Boston (1997)



Study on Optimal Model of Traffic Signal Control at Oversaturated Intersection

Chaoyang Wu^(✉), Xiaoqing Zeng, Jifei Zhan, and Qipeng Xiong

The Key Laboratory of Road and Traffic Engineering,
Ministry of Education, China, School of Transportation Engineering,
Tongji University, 4800 Cao'an Road, Shanghai, China
wcy@tongji.edu.cn

Abstract. With the rapid growth of automobiles in China, urban traffic congestion has become the most headache problem for citizens. In this paper, based on the SCATS of Shanghai urban traffic control, the traffic of oversaturated intersection is modeled and the optimization algorithm of intersection signal system is studied. The problem of traffic congestion can be alleviated by reducing the delay time and increasing the travel speed of vehicles at oversaturated intersections.

Keywords: Traffic congestion · SCATS · Oversaturated intersection · Optimization algorithm

1 Research Status

According to data released by the National Bureau of Statistics, in 2017, the total number of civil vehicles in China was 217.43 million (including 8.2 million tricycles and low-speed trucks), an increase of 11.8% over the same period last year, and a growth rate of 336.17% over 2006. Among them, the number of private cars was 18.69 million, an increase of 12.9%. Civil car ownership was 12.185 million, up 12.0%, including 11.416 million private cars, up 12.5%. With the increasing number of cars, traffic congestion has become increasingly prominent.

SCATS (Sydney Traffic Adaptive Coordination System) was developed by the New South Wales Road and Transportation Authority of Australia in the late 1970s. It has been widely used in Sydney and controls more than 3000 intersections. Later, it was installed and used in other cities. At present, about 50 cities in the world are running SCATS. Since 1985, Shanghai first introduced SCATS in China, and now more than 1300 intersections have been installed. Subsequently, Tianjin, Guangzhou, Shenyang, Hangzhou, Suzhou, Yichang, Ningbo and other cities began to use SCATS to control urban traffic. At present, about 50 cities in the world are running SCATS.

At present, the urban road traffic signal control system in Shanghai adopts SCATS. The data acquired by SCATS detection equipment and the demand of road traffic information collection in central city provide data basis for real-time release of traffic status, daily traffic management, traffic law enforcement and traffic safety management. SCATS obtains real-time road traffic status information through real-time processing of

extracted traffic information, and stores extracted traffic information and records historical traffic information. This paper is based on the SCATS of urban traffic control in Shanghai.

2 Research on Traffic State Perception Model of Intersection

In this paper, traffic state perception model is defined by traffic state partition of single entrance road and intersection. The traffic state division of single-entry road is based on queue length data, and the traffic state definition of intersection is based on the traffic state division of single-entry road.

2.1 Classification of Traffic Conditions of Single Entrance Road

The traffic state of a single entrance road can be defined by a two-point-three state model according to the green-light time at the intersection and the green-light time at the upstream intersection. The two points in the model refer to the saturation point and the overflow point, which are the two key points to determine the green light time from upstream to downstream.

Among them, the saturation point is the length of vehicle occupancy that can dissipate during the effective green light period of the intersection; the spillover point is the length that the intersection can accept the maximum number of vehicles in the upstream green light time and ensure that the intersection entrance queue will not spill on the intersection. The location of saturation point and overflow point is shown in Fig. 1.

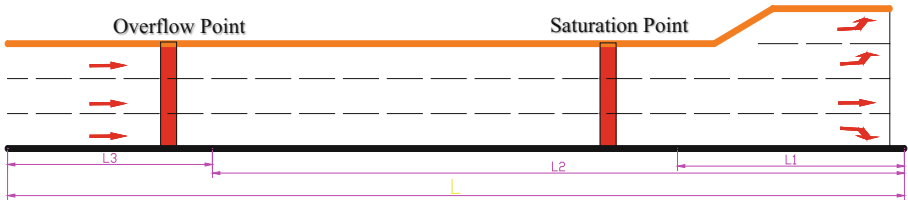


Fig. 1. Key point diagram

Saturation point is based on the matching principle of traffic capacity of intersection entrance and exit lanes. Calculating the queue length L_1 of all vehicles that can be cleared at an entrance during the green light period is the saturation point position. Taking the straight lane as an example, the defining formulas are as follows:

$$L_1 = \frac{S_1 \times f_1 \times g_e \times L_c}{3600 \times m_1} \quad (1)$$

Among them, L_1 : the position from saturation point to parking line, S_1 : the saturated flow of section, f_1 : the reduction factor of direct vehicles affected by steering vehicles at this intersection, L_c : body length, m_1 : number of lanes, g_e : the effective green light time at the intersection.

The spillover point considers the impact of queuing spillover on upstream intersections, calculates the number of vehicles released during the maximum green light time at upstream intersections, and determines the location of spillover point by calculating the threshold value of queuing length during congestion spillover. L_2 is the number of vehicles released during the maximum green light time at upstream intersections. Similarly, the straight lane is taken as an example.

$$L_2 = L - \frac{S_2 \times f_2 \times g_{max} \times Lc}{3600 \times m_2} \tag{2}$$

Among them: L_2 : the position from spillover point to parking line, S_2 : saturated flow of section, f_2 : the reduction factor of direct vehicles affected by steering vehicles at upstream intersection, m_2 : the number of lanes at upstream intersection, g_{max} : the maximum green light time at upstream intersection.

The three states in the two-point three-state model are determined by two key points. The queue length of intersection is defined as L_q . According to the saturation point and spillover point determined by the above principles, the intersection entrance can be divided into three states, which are as follows:

State 1: Unsaturated State

When the queue length is before the saturation point, that is, $L_q < L_1$, this state is called unsaturated state (state 1), as shown in Fig. 2. The queue length of the state can dissipate in a cycle of green light time under the original timing condition, so the original signal timing scheme meets the demand.



Fig. 2. Schematic diagram of unsaturated state (state 1)

State 2: Saturated Non-overflow State

When the queue length is between the saturation point and the overflow point, that is, $L_1 < L_q < L_2$, this state is called saturated non-overflow state (state 2), as shown in Fig. 3. The queue length of this state cannot dissipate in a green light period of a cycle under the original timing condition, so it needs to be adjusted.

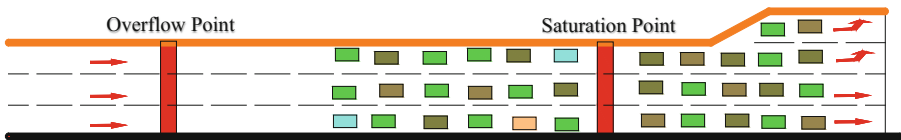


Fig. 3. Schematic diagram of saturated non-overflow state (state 2)

State 3: Saturation Overflow State

When the queue length exceeds the location of the overflow point, that is, $L_q > L_2$, this state becomes saturated overflow state (state 3), as shown in Fig. 4. The queuing length of the state will have a great impact on the upstream intersection, so the system needs to be fully coordinated in the second state to avoid the emergence of the state triple saturation overflow state.

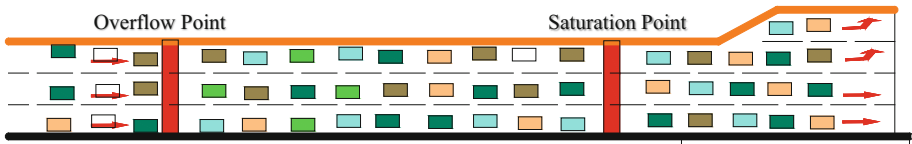


Fig. 4. Schematic diagram of saturation overflow state (state 3)

2.2 Definition of Traffic State at Intersections

In this chapter, when the queue length L_q exceeds the saturation point L_1 , it is saturated. The unsaturated state means that the queue length in all directions of the intersection does not reach the saturation point L_1 . Take a typical four-phase intersection (including East-West direct phase, East-West left-turn phase, north-south direct phase, north-south left-turn phase) as an example. Definition of traffic state at intersections is shown below.

State 1: Partial saturation state: queue length of partial entrance ($k < 4$) at intersection, $q > L_1$. As shown in Fig. 5.

State 2: Saturated non-spillover state: the queuing length of the four entrance lanes at the intersection is $q > L_1$, but the queuing length of all entrance lanes is $q < L_2$. As shown in Fig. 5.

State 3: Fully saturated spillover state: queue length of some entrances at intersections, $q > L_2$. As shown in Fig. 5.

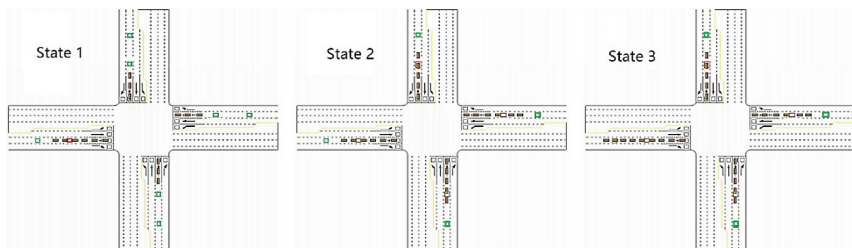


Fig. 5. Traffic state at intersections

In view of the above different intersection traffic conditions, different signal control strategies are adopted. The specific control method is shown in the next section.

3 Research on Optimal Control Algorithms of Signal System at Intersections

The principle of the optimization algorithm is to extend the green light time when the queue of a vehicle with an optimized phase (assuming that the phase is a north-south direct phase) reaches a certain length, so the phase applies for an extension of the green light time. However, the constraints should be taken into account if the green light extension application can be executed at this time. For example, it can be judged whether the restriction conditions such as whether the exit passage and other phases (e.g. lateral roads) are unobstructed or not are satisfied. When the exit or other phases (lateral roads) are congested, the signal scheme is not adjusted. Optimal Control Algorithms for Signal Systems at Intersections is shown below.

3.1 Model Index and Calculation of Optimization Algorithm

The optimization index F in the optimization model consists of three parts: the queue length index f_q of the entrance lane, the speed f_v of the exit lane and the conflict phase congestion state f_n .

The queue length index f_q of the entrance lane ranges from $\{0, 1, 2, 3\}$ to represent the four states of queue length, as follows:

$$f(L_q) = \begin{cases} 0, L_q < q_0 \\ 1, q_0 < L_q < q_1 \\ 2, q_1 < L_q < q_2 \\ 3, q_2 < L_q < q_3 \end{cases} \quad (3)$$

The corresponding situation can be described in Fig. 6.

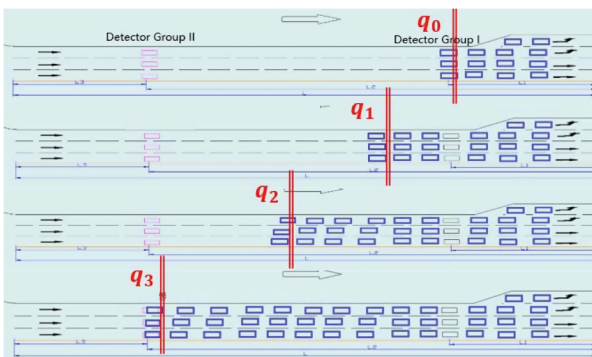


Fig. 6. Schematic diagram of queuing indicators for inlet lanes

When the queue length reaches the above target, the phase can request the green light time extension. Whether the request can be allowed by the system also needs to be determined by other factors (such as exit lane speed, conflict phase congestion, etc.).

The allowable green light extension time of the system consists of three levels, corresponding to the extension of t_1, t_2, t_3 , ($t_1 < t_2 < t_3$), the green light extension t_1 can be proposed when $f(L_q) = 1$, the green light extension t_2 can be proposed when $f(L_q) = 2$, and the green light extension t_3 can be proposed when $f(L_q) = 3$. Especially, when $f(L_q) = 0$, the green light extension t_3 can meet the requirements without lifting. Extension request.

The position of q_0 is the maximum queue length of the green light time in a cycle, which is the same as the saturation point of the single inlet L_1 mentioned above. This shows that in the above case, when $f(L_q) = 0$, there is no need to ask for the green light time extension. The maximum queuing length in the location optimization interval of q_3 is the same as that of the single inlet spillover point L_2 mentioned above.

Accordingly, the positions of q_1 and q_2 are determined as follows:

$$q_1 = L_1 + (L_2 - L_1) \times \frac{t_1}{t_3} \quad (4)$$

$$q_2 = L_1 + (L_2 - L_1) \times \frac{t_2}{t_3} \quad (5)$$

The queuing length index f_v of exit corridor ranges from 0 to 1 to indicate the congestion degree of exit corresponds to the congestion and unobstructed situation of exit corridor. Definition:

$$f_v(V) = \begin{cases} 0, & V < 5 \text{ km/h} \\ 1, & V \geq 5 \text{ km/h} \end{cases} \quad (6)$$

The conflict phase congestion index f_n is defined as $\{0, 1\}$ to represent the degree of phase congestion that is in conflict with the phase of the optimization request. It is expressed by the queue length of the conflict phase and corresponds to the phase supersaturation and unsaturation. Definition:

$$f_n(N) = \begin{cases} 0, & L_q \geq L_1 \\ 1, & L_q < L_1 \end{cases} \quad (7)$$

Then the optimization index of the phase is:

$$F = f_q * f_v * f_n \quad (8)$$

3.2 Selection of Optimal Control Scheme

Comparing the optimization index between two phases $F = f_q * f_v * f_n$, the corresponding optimization scheme is selected according to the comparison of tables. The comparison table is shown in Table 1.

Table 1. Optimized schemes selection table

Optimization index		Phase I			
		0	1	2	3
Phase II	0	–	I, t1	I, t2	I, t3
	1	II, t1	I, t1	I, t2	I, t3
	2	II, t2	II, t1	I, t1	I, t2
	3	II, t2	II, t2	II, t1	I, t1

3.3 The Relationship Between Intersection State Definition and Optimal Control Algorithm

State 1: Optimizing strategy under partial saturation state

When F_n ($n = 1, 2, 3, 4$) has any item greater than 0, but the queue length $q < L_2$ of all entrance lanes can reduce the green light time of non-crowded phase and compensate it to the crowded phase, so that the crowded phase has more traffic time and more traffic rights, so as to alleviate the traffic congestion in the crowded phase.

State 2: Optimizing strategy under full saturation and Non-overflow state

When F_n is more than 1, but queue length $q < L_2$ for all entrance lanes. In this traffic state, because the queue length in all directions is very large, the principle of minimum total queue length at intersections should be taken, and different weights should be given according to the flow of different phases, so as to select the signal optimization scheme.

State 3: Optimizing strategy under fully saturated spillover

F_n is all greater than 0, ($n = 1, 2, 3, 4$) and there is any inlet $q > L_2$. In this traffic state, because queue length in all directions is very large and may affect the upstream intersection, it is necessary to reduce the number of vehicles entering the bottleneck phase of the bottleneck intersection in the congestion time, that is, to reduce the green light time at the upstream intersection, so as to reduce the queue length in this direction.

4 Simulation

4.1 Simulation Data Acquisition

Taking Huaihai Road-Chongqing Road intersection in Huangpu District of Shanghai as the simulation object, this paper conducts on-site traffic investigation on the intersection. According to the actual data, the dynamic control strategy based on queue length detection proposed in Sect. 4 is optimized on the basis of the existing signal timing model. The simulation model is established by using the micro-simulation software

VISSIM, and the simulation is carried out through the COM interface of VISSIM. Secondly, the online simulation of the control strategy is realized, and the effect of the optimal control strategy is evaluated.

Choose the working day, morning and evening peak time to conduct traffic survey at intersections. After traffic survey, the basic situation of the simulation intersection is sorted out as shown in Table 2.

Table 2. Simulated intersection basic situation questionnaire

Import Road	Lane function		Number of lanes	Early peak flow (pcu/h)		Late peak flow (pcu/h)		Green time(s)	Distance from adjacent intersections (m)
North import	Straight line		4	255	1307	321	1374	60	125 (315)
				517		469			
				104		175			
				431		409			
South import	Ground	Left turn	1	269	269	170	170	24	305
		Straight line	2	527	1047	313	640	60	
	Ramp	Left Turn	1	290		290	204	204	
		Straight line	1	694	694	579	579	84	
West import	Left turn		1	224	224	224	224	35	320
	Straight line		1	399	399	431	431	51	
East import	Left turn		1	209	209	280	280	35	68 (180)
	Straight line		1	230	230	239	239	51	

According to the survey data, the traffic volume in the north-south direction is larger during the morning and evening rush hours, and the difference between the East and the west direction is larger. Especially, influenced by the direction of residents' travel to work, during the early rush hours, the ground and elevated traffic flow from south to north are larger, and ramp traffic jams to elevated lanes frequently occur. Considering that the east-west traffic flow is less and there is certain optimization space, the North-South direct phase is chosen as the main phase to optimize, and the real traffic data of the early peak of the intersection is used as the input of the simulation to verify the optimization control strategy.

The current signal timing scheme has a period of 170 s and consists of five phases. The green light duration and direction of each phase are shown in Fig. 7.

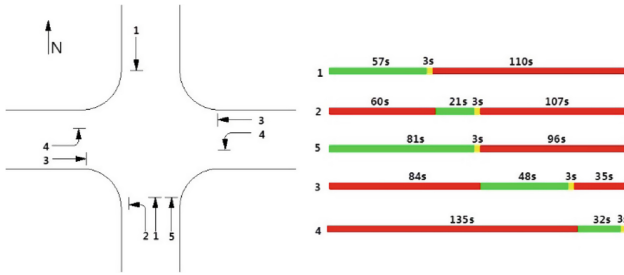


Fig. 7. Simulated signal timing diagram of intersection

4.2 Simulation Modeling Process

Through VISSIM simulation software, the simulation model is established according to the actual situation. The detector layout is shown in Fig. 8. In order to make the simulation reflect the real situation of the intersection and the congestion characteristics of the bottleneck intersection, the upstream and downstream intersections related to it are integrated into the road network.

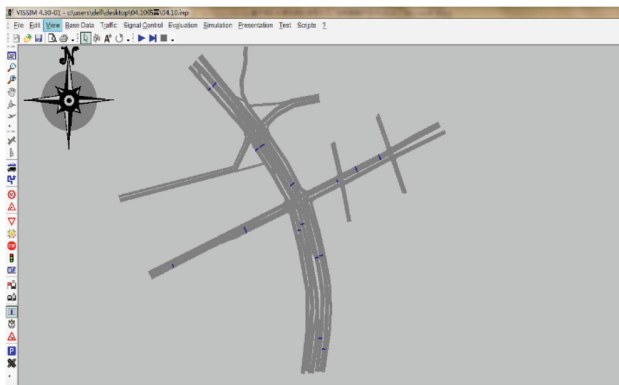


Fig. 8. Schematic diagram of simulated basic road network and detector layout

Considering the special situation of the simulation intersection, there are short-line intersections in some directions. In this case, this study considers that the queue detector I can be cancelled and only the overflow detector can be installed, that is to say, the prevention of vehicle queue overflow is the primary goal. The locations of detectors in each inlet are shown in Table 3.

In order to implement the signal control algorithm based on queue length detection proposed in this paper, the C# programming language is used to program the algorithm. C# is an object-oriented programming, encapsulating the code with unique properties, making the code safer and enhancing the maintainability and expansibility.

Table 3. Location of inlet channel detectors

Detector location (distance from parking line, unit: m)	North	South				West		East	
	Straight line	Ground		Ramp		Straight line	Left Turn	Straight line	Left Turn
		Straight line	Left Turn	Straight line	Left Turn				
Detector I	110	110	45	110	45	90	75	90	90
Detector II	265	255	105	300	300	255	255	160	160

4.3 Analysis of Simulation Results

In this study, queue length, travel time and delay are used to evaluate the optimization algorithm proposed above. The simulation under the optimization control algorithm is carried out three times and compared with the current scheme. The following three indicators are analyzed separately in Table 4.

Table 4. Comparison of simulation results before and after optimization

Traffic signal control indicators	Simulation serial number	South entrance straight lane			
		Ground	Contrast (%)	Ramp	Contrast (%)
Average queue length	Before optimization	167.86	–	175.73	–
	Sim1	161.41	–3.8	168.47	–4.1
	Sim2	160.36	–4.5	165.82	–5.6
	Sim3	160.77	–4.2	164.9	–6.2
Travel time	Before optimization	111.56	–	75.90	–
	Sim1	107.38	–4.1	71.44	–5.9
	Sim2	100.63	–10.1	71.17	–6.2
	Sim3	102.16	–8.8	70.63	–6.9
Average travel speed	Before optimization	13.1	–	19.3	–
	Sim1	13.6	3.9	20.5	6.0
	Sim2	14.5	10.9	20.5	6.4
	Sim3	14.3	9.2	20.7	7.2
Vehicle travel delays	Before optimization	77.2	–	40.9	–
	Sim1	73	–5.4	36.4	–11.0
	Sim2	66.3	–14.1	36.2	–11.5
	Sim3	67.9	–12.0	35.6	–13.0
Vehicle Parking Delay	Before optimization	57.7	–	29.5	–
	Sim1	53.8	–6.8	25.6	–13.2
	Sim2	48.2	–16.5	25.2	–14.6
	Sim3	49.8	–13.7	25.4	–13.9

(continued)

Table 4. (continued)

Traffic signal control indicators	Simulation serial number	South entrance straight lane			
		Ground	Contrast (%)	Ramp	Contrast (%)
Traffic signal control indicators	Before optimization	Travel delay	Contrast (%)	Parking delay	Contrast (%)
Average delay at intersection	Sim1	47.3	–	36.6	–
	Sim2	43.7	–7.6	33.1	–9.6
	Sim3	41.2	–12.9	31.2	–14.8
	Before optimization	41.3	–12.7	31.5	–13.9

5 Summary

In this paper, the traffic signal control optimization model of oversaturated intersection is mainly studied. This paper uses two-point-three state traffic state perception model to analyze the traffic state. Finally, the optimization system of intersection signal system is validated and evaluated by simulation. The optimization results can increase the vehicle speed by more than 5% in the peak period and the overall average speed by more than 10%.

Acknowledgments. The project is supported by Key Project of Special Development Fund for Zhangjiang National Independent Innovation Demonstration Zone in Shanghai (Number 207105-JD-C1085-072, 201705-JD-C1085-056). This project is supported by the Shanghai Science and Technology Committee (Grant No. 17DZ1204003).

References

1. Zeng, X., Wu, C., Chen, Y., et al.: Research on the model of traffic signal control and signal coordinated control. In: Proceedings of the International Symposium for Intelligent Transportation and Smart City. Springer. EI: 20171603584791 (2017)
2. Zeng, X., Zhan, J, Yang, L., et al.: Research on the queue length prediction model with consideration for stochastic fluid. In: Proceedings of the International Symposium for Intelligent Transportation and Smart City. Springer. EI: 20171603584790 (2017)
3. Akcelik, R.: Traffic signals: capacity and timing analysis. ARRB, Research Report 123, Victoria, Australia (1981)
4. Wang, Y., Jin, L., Yuan, T.: Research on traffic flow state theory of vehicle queuing at Urban Road intersection. Urban Construction Theory Research: Electronic Version, 1 (2016)
5. Yang, L., Zeng, X.: The research of signal optimal control strategy based on SCATS. Appl. Mech. Mater. **4**, 196–202 (2015)
6. Cipriani, E., Fusco, G.: Solution procedures for the global optimization of signal settings and traffic assignment combined problem. In: Proceedings of ICTTS, Traffic and Transportation Studies, pp. 325–336 (2002)



Application of Intelligent Transportation System in Intelligent Network Environment

Qiulan Wang^(✉) and Sayi Wang

Shanghai Urban Construction Design and Research Institute (Group) Co., Ltd.,
Shanghai, China
wangqiulan@sudri.com

Abstract. With the advent of global information and technological revolution, Intelligent Connected Vehicle (ICV) has become an emerging industry. It is the hotspot for global innovation and the critical element to driving future development. This paper is dedicated to an analysis of the key techniques of ICV applied in the automated driving environment. Along with the applications of many intelligent transportation application systems. With a combination of Traffic Sign Control system, Pedestrian Detection system, Electronic License Plate system, Smart Parking system, Traffic Operation Status Analysis system and Traffic Accident Warning system, ICV is developed with the capability of collecting and controlling the environmental information. Therefore, it has facilitated the development of intelligent network industry, leading to the improvement in traffic safety and efficiency for urban regulations, as well as the levels of satisfaction with daily travels among the general public.

Keywords: Intelligent network · Automated driving · Intelligent transportation · On-road vehicle road coordination

1 Overview

Intelligent and networked car represents an emerging industry against the backdrop of the revolution in information technology. It has come under the spotlight for innovation across the world and become a critical factor in driving future development. The advanced in vehicle sensors, controllers, actuators and other devices are representative of what combines modern communication and network technologies to deliver the features of V2X intelligent information exchanging and sharing of X (people, vehicle, roads, backgrounds, etc.). The new generation of the vehicle's control system is intended to replace human as the drivers [1, 13]. Intelligent networked vehicles is designed to provide safer, more energy-efficient, more eco-friendly and more convenient ways of travel. Therefore, it has a promising prospect of leading the global automotive industry to transform the landscape of technological advancement and future development, which constitutes an important part of the intelligent transportation system in urban development.

Intelligent networked vehicle is much more than providing a means of travel, which is because it involve a range of major technologies as required to operate it, such as environmental awareness, intelligent decision-making, and coordinated controlling.

Besides, it requires various information interactions between one vehicle and another, vehicles and roads, as well as vehicles and people [2–4]. At present, developing intelligent network connection technology is largely reliant on the sensors and algorithms relating to them. If it could be made possible that information exchanges between the intelligent network connected with vehicles and the intelligent transportation system, a multi-sourced, multi-dimensional, dynamic and intelligent communication system will become a reality. Therefore, by promoting the development of intelligent network connection industry, and simultaneously carrying out the deployment of intelligent vehicle road coordination and cloud-based traffic control system, a comprehensive improvement will be made in urban traffic for its safety and efficiency, as well as in the satisfaction levels with public transport among citizens and efficiency of traffic regulations.

2 The Key Technology of Intelligent Network

The key technologies applied to networked vehicles running on the road mainly include environmental awareness technology, independent decision-making technology and control execution technology [5, 6].



(1) The technology of environmental awareness

The driving of intelligent networked vehicle first requires to collect and process various environmental information and in-vehicle information. The surrounding environmental information required by the vehicle mainly includes road boundaries, nearby vehicles, pedestrians and so on. Currently, as far as a majority of the environment awareness technology solutions are concerned, it relies largely on various equipments installed on the vehicle, such as laser radar, camera, millimeter wave radar, GPS positioning device, inertial navigation device, and other forms of ultrasonic sensors, infrared sensors and other sensors.

(2) The decision-making technology

Once the intelligent networked vehicle obtains real-time environmental information from the surrounding area, it needs to make decision and judgment, before determining the most appropriate working model and deciding the corresponding control strategies, such as lane keeping, lane departure warning, distance keeping, obstacle warning, etc. It is also required to predict the status of vehicles and other vehicles, lanes, pedestrians, etc. in the future.

(3) The control execution technology

Control execution means that the intelligent vehicle needs to find a passable path according to a certain searching algorithm, thereby realizing autonomous navigation of the intelligent vehicle based on the perception of environmental information and determining the position of the vehicle in the environment. According to the planned driving route, the intelligent networked vehicle is capable of travelling along the desired trajectory.

At present, the downside of intelligent networked car is that the automated driving is excessively reliant on weather conditions, terrains and the quality of map. In addition, it lacks the understanding of unknown environmental conditions of some special events. For the complicated traffic conditions, the unmanned vehicle would fail to cope. In order to meet the demand for intelligent networked automatic driving, it is necessary to rely on the intelligent transportation system to deliver the knowable, controllable and serviceable functions, and to provide support for the safe operation of intelligent networked vehicles.

3 Intelligent Transportation System Application in Intelligent Network Environment

3.1 Smart Road Network

The smart road network mainly enables the correct identification of surrounding environment, which is crucial to improving the safety of vehicles, pedestrians and the environment, and to ensuring that intelligent networked vehicles would be able to operate safely and effectively on the road to provide driving environment for intelligent networked vehicles.

(1) The intelligent traffic signal controlling system based on V2X

The V2X intelligent traffic signal controlling system is applied mainly to read signals through the intelligent road terminal (Sibox) located on the roadside, and then transmit it to the vehicle connected intelligent network. With the advancement of technology, the intelligent roadside terminal can be incorporated into the signal machine. Thus, the signal machine actively sends information to the intelligent networked vehicle [7] (Fig. 1).

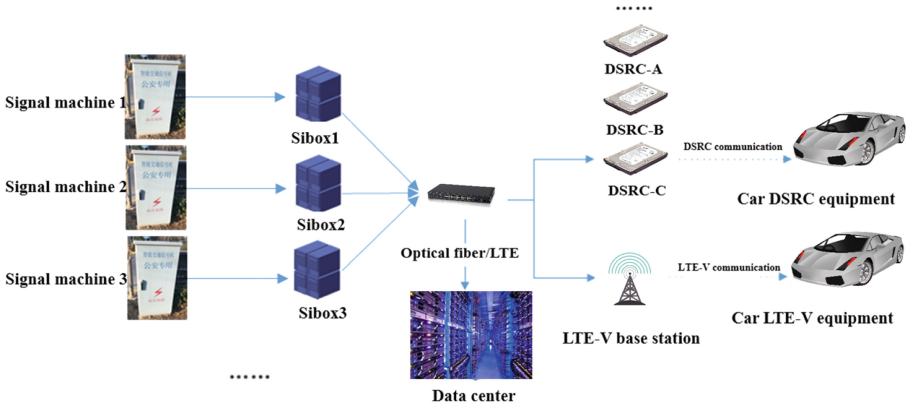


Fig. 1. The framework diagram of the intelligent traffic signal control system based on V2X

The core equipment included in the V2X intelligent traffic signal control system is the roadside intelligent terminal equipment (Sibox), which is deliberately placed beside the existing signal control machine to allow for the intelligent transformation of the signal control intersection, provide an open signal control data interface, and release the signal in real time. Information is needed to meet the vehicle road communication (V2I) test and application. Its specific features include:

- (1) Reading the signal data from the signal communication port (network port or serial port);
 - (2) Reading the signal data from the lamp controlling output terminal by current and voltage sampling;
 - (3) The protocol conversion function;
 - (4) The function of the vehicle-mounted green wave speed guidance and signal priority interface;
 - (5) The traffic flow detecting device;
 - (6) Local storage configuration files, fault logs, and operating records;
 - (7) Communicating with the center to sending device status and relevant data.
- (2) The pedestrian detection system based on V2X

Pedestrian detectors are usually installed on the roadsides of intersections and corners to collect the real-time information of pedestrians showing up in the roads in an all-dimensional way, analyze and assess the routes of pedestrians, and issue warnings of blind spots in the line of sight through the network interconnection to avoid collision, which makes both pedestrians and vehicles effectively coordinated with roads. With the assistance of these devices, improvement can be made in traffic safety and traffic accidents will be significantly reduced as a result.

The system is mainly comprised of three parts: the front-end pedestrian detector, the radar system, and the pedestrian detection controlling box.

Pedestrian detectors are installed at a height of 4–6 m clear of the ground. Each detector is able to detect a minimum of 3 lane-wide pedestrians crossing the road through a pre-setting presence detection zone [8]. The detection information/video is then transmitted to the pedestrian detection controlling box for subsequent processing by connecting the cable (Fig. 2).

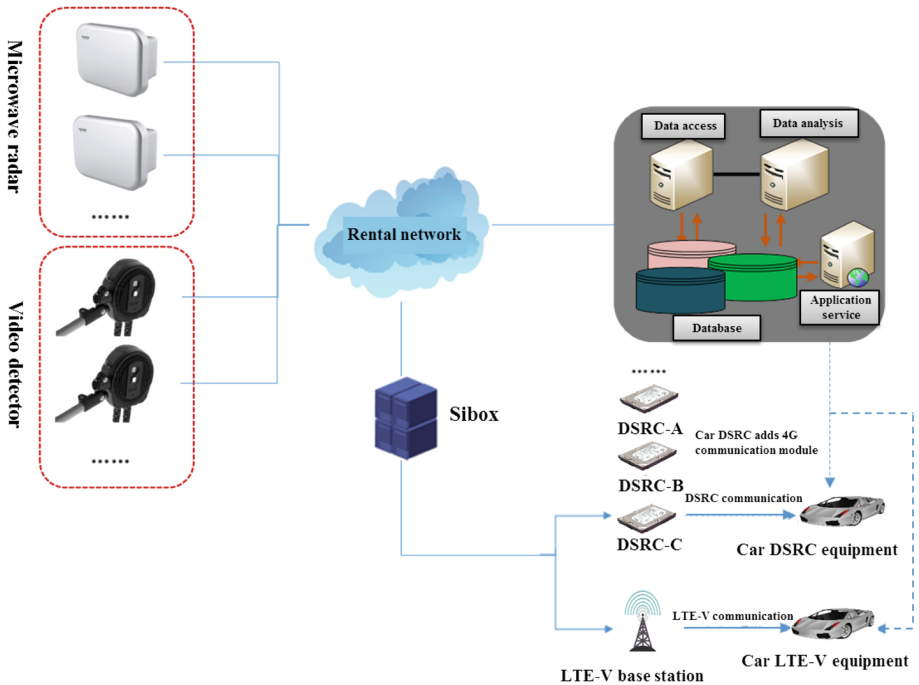


Fig. 2. The framework diagram of the pedestrian detection system based on V2X

(3) Vehicle detection and application system based on automobile electronic identification

The automobile electronic identification is purposed to achieve the refined collection of intelligent networked car information. With this system, the level of vehicle management information can be raised to accomplish the application demonstration of automotive electronic identification technology in the field of intelligent networked vehicles.

An access to the vehicle electronic identification data and management function module are constructed in the intelligent networked automobile integrated data center, which provides an access to the data collected by the automobile electronic identification system in the open road testing area of the intelligent network-connected core area. The data is stored, managed and analyzed, and the refined intelligent networked vehicle information (inclusive of the car engine number, vehicle model, annual review

status, etc.) is collected to enable the improved management of the information on intelligent networked vehicles (Fig. 3).

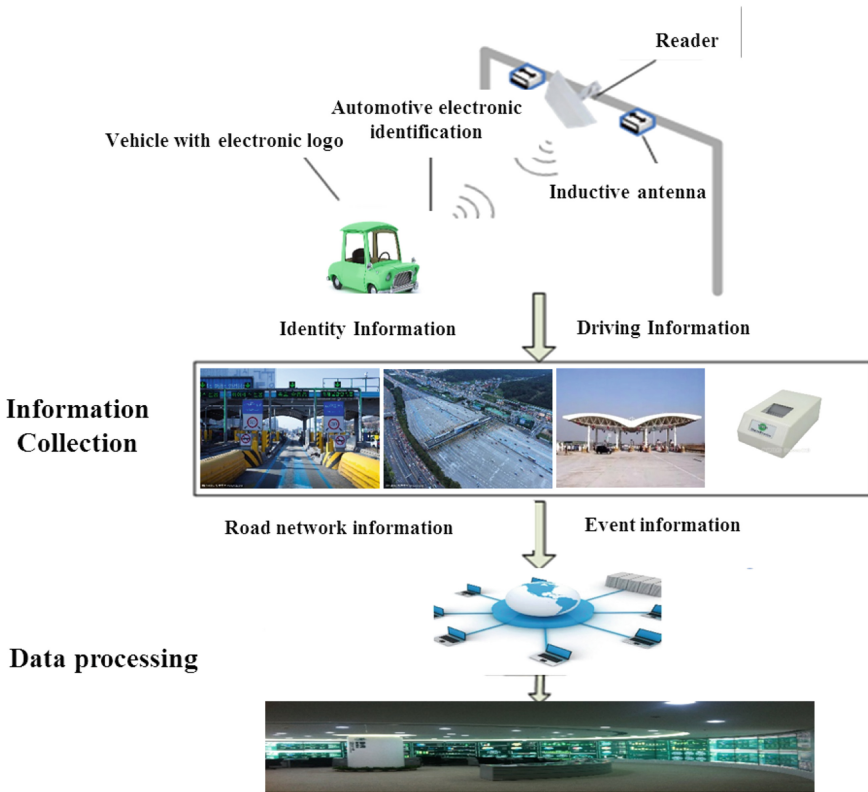


Fig. 3. The framework diagram of the automotive electronic identification system

The construction of vehicle detection and application system is based on the automobile electronic identification system, which mainly requires the automobile electronic identification reader with antenna layout; the transformation of customized bus and the installation of the automobile electronic identification system; the demonstration of the priority application of the customized bus signal based on the automobile electronic identification technology, data collection and urban traffic information management supporting.

3.2 Automatic Parking and Wireless Charging

Parking spaces in large cities are usually limited. It has become a required skill to drive cars into tight spaces. At the same time, parking is difficult, which may result in traffic congestion, nervous exhaustion and minor collisions.

The intelligent networked vehicle can facilitate automatic parking and wireless charging, which maximizes the usable space and enables intelligent management, while saving user’s time and making the city more humanized.

By utilizing radar, video surveillance and other sensing technologies, the holographic perception of the parking lot and surrounding environment is enabled. Through the sensor, plenty of information including the signals of the surrounding vehicles, the size of the parking space and the distance from the roadside are all transmitted to the intelligent networked vehicle, with which the vehicle is guided into the parking space smoothly through the decision-making technology and the control execution technology (Fig. 4).

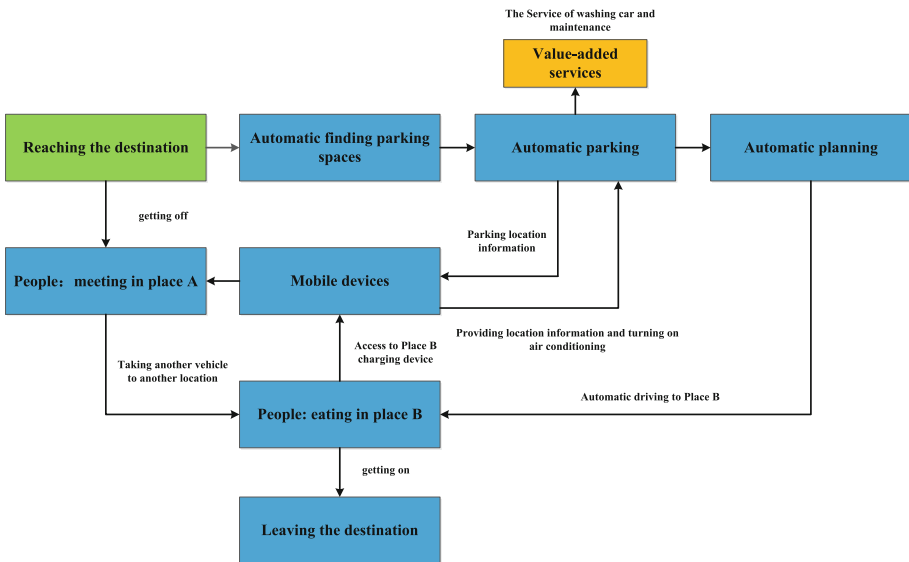


Fig. 4. The framework diagram of automatic parking

3.3 The System of Vehicle-Road Coordination Application

(1) Triggering the traffic status information

Balancing the distribution of road network traffic flow is an ideal scenario for urban road traffic operations [7]. Traffic managers would like to guide the driver to opt for the right path to travel through reasonable traffic information. From a macro perspective, the goal of equal distribution of traffic flow on the road network can be achieved; from the micro view, the delay to the driver’s travel can be reduced. The issue of real-time dynamic traffic information is a required measure for dynamic path guidance, which is understood to be convenient for the drivers to make the right and sensible route decisions based on the published information.

In order to ensure that the entering and exiting vehicles can find a reasonable route to reach their destination, and avoid the most congested road sections and the road sections where traffic accidents occur the intelligent transportation system uses a combination of big data, Internet and other high-tech approaches to analyze and mine the data relating to traffic operation status, and publish the macro, easy-to-understand traffic indicators for all of the road users.

The multi-mode information communication system is based on intelligent network connection vehicle V2X [9], in order to improve the operation efficiency of intelligent vehicles and perceive the traffic operation status in advance, the multi-mode information communication system can be adopted to realize the interconnection and communication between the traffic status information and the intelligent network connection vehicle, as a result, the vehicle can know about the real-time road condition in a timely manner, and also know various information, such as traffic propaganda and traffic notices, in order to further plan their appropriate route for travel, reduce traffic congestion and minimize travel delays.

(2) Traffic event forecasting and warning

The traffic event forecasting and warning system is used to integrate all the information on vehicles and roads into the real-time interactive V2X multi-mode communication network. These cars can perceive the surrounding vehicles and road information before making reasonable driving behavior in advance and path planning; at the same time, through coordination with big data processing and according to the control strategy, real-time controlling or induction schemes for vehicles, signal lights, etc., road operation is made safe and smooth.

(1) The warning of the road water-logging

With the application of existing in-vehicle equipments, roadside equipments and dedicated short-range communication wireless module, the liquid level information is broadcasted to the intelligent networked vehicle with high-speed movement within 2 km. Based on the current state of the vehicle and the position of the water accumulation point, the hierarchical safety distance assessment model delivers the liquid level and alarming information to the intelligent networked vehicle either through WIFI or Bluetooth module, with which the vehicle can make response by taking measures such as slowing down or bypassing to drive.

(2) The warnings of the road construction

Adding the existing road construction signage into the high-precision map [10–12], which includes construction location information, construction lanes, lane-changing reminder information, construction type, construction time and other information. The intelligent network vehicle is capable of selecting the most appropriate path based on the information. In the meantime, the information is transmitted back to the command center through the 4G module, which could facilitate the real-time monitoring and the assessment of traffic impacts caused by road constructions.

(3) The warnings of the traffic accident

Based on the analysis system of traffic big data, using a combination of intelligent video processing, neural network, association rules and other technologies which add up and forecasts traffic data of those designated road sections. Using video detection technology combined with weather, road section conditions (such as number of lanes) and other information, the hazard conditions of road section is more precisely assessed, accident predictions are made more accurately, then an alarm is issued. Through V2X multi-mode communication technology, the warning information about traffic accident prediction is sent out to the intelligent networked vehicle, based on which some countermeasures can be taken in advance to lower the probability of traffic accidents occurring.

4 Conclusion

At the present time, the realization of autonomous driving is mainly dependent on the environmental awareness system, decision-making algorithm and controlling execution technology of the intelligent networked car. Nonetheless, it is lacking in the understanding and response ability for the unknown environments relating to some special events. Therefore, the application of intelligent transportation system in the intelligent networked environment is carefully analyzed by this paper through studying the key technologies of intelligent driving in the intelligent networked car. The perception and controlling ability of intelligent networked vehicles on environmental information are further refined, which is believed to be conducive to promoting the development of the intelligent network industry, comprehensively improving traffic safety and traffic efficiency in urban areas, as well as public travel satisfaction and the efficiency in urban traffic regulations.


References

1. Fu-quan, P., Rong-jie, Z., Wei, Z., Li-xia, Z.: Research and development prospects of driverless vehicles. *Sci. Technol. Innov. Appl.* **2**, 27–28 (2017)
2. Thrun, S.: Toward robotic cars. *Commun. ACM* **53**, 99–106 (2010)
3. Yang, Z.: Key Technologies of Driverless Vehicles. *China Sci. Technol. Expo* **26**, 272 (2011)
4. Qiao, Y.: Driverless cars open a new era of “intelligent traffic”. *Second Classroom* **11**, 4–9 (2014)
5. Min, Y.: Research status and development direction of driverless vehicles. *Automob. Maint.* **02**, 9–10 (2003)
6. Xueqiang, F., Liangxu, Z., Zhizong, L.: Overview of the development of driverless cars. *Shandong Ind. Technol.* **05**, 126 (2015)
7. Winston, C., Mannering, F.: Implementing technology to improve public highway performance: a leapfrog technology from the private sector is going to be necessary. *Econ. Transp.* **3**(2), 158–165 (2014)
8. Xueqiang, F., Liangxu, Z., Zhizong, L.: Overview of the development of driverless cars. *Shandong Ind. Technol.* **5**, 51 (2015)

9. Lijing: Google driverless car: from transformation to self-made. Beijing Science and Technology News, 2014-06-09034
10. Huang, W.: Whether a driverless car make the city unblocked. Guangming Daily, 2012-06-09034
11. Ma, T.: Research on the motion stability of driverless smart cars. XiAn Industrial University (2013)
12. Jianliang, P.: Analysis of social benefits and impacts of driverless cars. Autom. Ind. Res. **5**, 22–24 (2014)
13. Bellamy, D., Pravica, L.: Assessing the impact of driverless haul trucks in Australian surface mining. Resour. Policy **36**(2), 149–158 (2011)



Stochastic Traffic-Assignment with Multi-modes Based on Bounded Rationality

Zhi Zuo¹ , Xiaofeng Pan², Lixiao Wang¹, and Tao Feng²

¹ School of Civil and Architectural Engineering,
Xinjiang University, Urumqi, China

zuo zhiukk@foxmail.com, xjwanglx@foxmail.com

² Urban Planning Group, Department of the Built Environment,
Eindhoven University of Technology, 5600MB Eindhoven, Netherlands
{x.Pan, t.feng}@tue.nl

Abstract. This paper proposes a stochastic traffic assignment model with multi-modes incorporating the concept of bounded rationality. Multi-criteria decision is considered using TODIM (which stands for “multi-criteria, interactive decision making” in Portuguese) method to generate variable demands, route uncertainty is taken into account based on cumulative prospect theory to measure route choice behavior. A numerical example is used to verify the validity of the new model. The sensitivity of the scaling parameters for the mode and route choice is also analyzed. Results confirmed the model’s applicability and showed that travelers’ preferences on different routes are reference dependent. Two scaling parameters have a significant influence on the final results and must be estimated very carefully from real data.

Keywords: Travel behavior · Bounded rationality · Variable demands · Cumulative prospect theory · TODIM method

1 Introduction

Discrete choice models have been widely used to analyze the choices that travelers make [1–4]. Many previous studies were based on the assumption of the absolute rationality of travelers, i.e., that decision makers have specific goals, perfect knowledge and consistent preferences [5]. However, an increasing number of studies have suggested that the rationality of travelers is bounded [6–8].

In the mid-twentieth century, researchers began to question the validity of the complete rationality assumption [9, 10]. Simon introduced bounded rationality theory, in which “the term bounded rationality is used to designate rational choice that takes into account the cognitive limitations of the decision maker – limitations of both knowledge and computational capacity” [11]. This theory assumes that choices are based not on optimization but rather on satisfaction, i.e., that decision makers will accept a satisfactory alternative over the optimal one. Numerous studies of the decision-making process have been conducted [12–15]. Although various theories have been proposed, prospect theory and cumulative prospect theory [16, 17] have received the most attention [18–20].

In transportation research, cumulative prospect theory (CPT) has often been used to predict travelers' choices. Senbil and Kitamura [21] used CPT to assess reference dependence in commuter departure time choices. Jou et al. [22] analyzed dynamic commuter departure time choices under conditions of uncertainty using CPT. Both studies defined two reference points, the earliest acceptable arrival time (or preferred arrival time) and the work start time. Liu and Lam [23] investigated the effect of population density on travelers' mode choices using CPT. Connors and Sumalee [24] studied travelers' route choices and proposed a network equilibrium model based on CPT. Xu et al. [25] proposed a prospect-based user equilibrium model with an endogenous reference point and tested its application to congestion pricing. The problem of calibrating the parameters of CPT models for transportation applications has been investigated. Avineri and Bovy [26] identified the parameters for a CPT travel choice model and proposed several approaches to determine the parameter values. Jou and Chen [27] analyzed drivers' attitudes and preferences regarding risk and estimated the parameters for a CPT model. They found that the behaviors of drivers in Taiwan could be adequately modeled with CPT if real-time traffic information was available.

Although a number of studies have used CPT to model travelers' choices, there are certain limitations to this theory [28–31]. One problem is its weakness in addressing multi-criteria decision making, especially when qualitative determinants are included. To address this problem, Gomes and Lima [32] introduced the TODIM method, which is based on CPT but without its disadvantages. After decades of development, a large number of applications of the TODIM method have appeared in various fields [33, 34]. However, to the best of the authors' knowledge there have been few applications of this approach in transportation research. Only the study by Quadros and Nassi used the TODIM method to perform multi-criteria analysis for transportation, specifically, transportation-infrastructure investments in Brazil.

Most approaches based on bounded rationality include a user equilibrium (UE) model or a stochastic user equilibrium (SUE) model. These models, known as assignment models with constant demands, do not consider the effects that certain choice dimensions such as trip production, departure time, destination, and transportation mode may have on the equilibrium configuration [35]. Assignment models with variable (elastic) demands are required because demand elasticity may be relevant for urban planning over a medium- to long-term horizon.

This study investigates travelers' route choices and proposes a new stochastic user equilibrium model with variable demands based on bounded rationality. The model considers the effects of transportation modes on traffic demand (volume) and combines mode splitting with traffic assignment. The normalized evaluation matrix of the TODIM method is used to compute the probability of choosing travel in a privately owned vehicle (i.e., driving). The levels of demand for travel by privately owned vehicles and other modes are obtained, and stochastic user equilibrium assignment is performed based on cumulative prospect theory.

The remainder of this paper is organized as follows. Section 2 describes the framework and the main concepts of CPT and the TODIM method. Section 3 presents the stochastic user equilibrium model with variable demands and a heuristic algorithm, and a numerical example is presented in Sect. 4. Section 5 provides a summary and conclusions.

2 Methodology

2.1 Cumulative Prospect Theory

Following Simon's introduction of bounded rationality, Kahneman and Tversky [16, 17] questioned the validity of expected utility theory (EUT) as a useful approach for modeling human decision making and suggested cumulative prospect theory as an alternative. Three principles, reference dependence, loss aversion and nonlinear probability weighting, distinguish CPT from EUT. A value function and a probability weighting function are used to define the prospect of each alternative:

$$V = \sum_{i=0}^n \pi_i^+ v(x_i) + \sum_{i=-m}^0 \pi_i^- v(x_i) \quad (1)$$

where

$$\pi_n^+ = w^+(p_n) \quad (2)$$

$$\pi_{-m}^- = w^-(p_{-m}) \quad (3)$$

$$\pi_i^+ = w^+(p_i + \dots + p_n) - w^+(p_{i+1} + \dots + p_n), \quad 0 \leq i \leq n-1 \quad (4)$$

$$\pi_i^- = w^-(p_{-m} + \dots + p_i) - w^-(p_{-m} + \dots + p_{i-1}), \quad 1-m \leq i \leq 0 \quad (5)$$

In (1)–(5), V is the prospect of an alternative, $v(\cdot)$ and $w(\cdot)$ are the value function and the probability weighting function, respectively, $\pi(\cdot)$ is the cumulative probability weighting function, x_i represents the difference between outcome i and the reference point, p_i denotes the probability of outcome i , and m and n are the numbers of negative and positive prospects, respectively.

The value function $v(\cdot)$ in CPT is an improvement over the utility function in EUT. According to Tversky and Kahneman [36], “the carriers of value are gains and losses defined relative to a reference point” and “the function is steeper in the negative domain than in the positive one”. The value function can be expressed as follows:

$$v(x_i) = \begin{cases} x_i^\alpha & \text{if } x_i \geq 0 \\ -\lambda(-x_i)^\beta & \text{if } x_i < 0 \end{cases} \quad (6)$$

where $\lambda > 1$ and α and β are positive parameters.

The probability weighting function $w(\cdot)$ in CPT can be treated as a regression of the objective probability. This function is nonlinear (reverse S-shaped) and monotonically increasing. For route choices, this function can be expressed as follows (Prelec 1998; Ingersoll 2008; Zuo et al. 2015) [37–39]:

$$w(p_i) = \begin{cases} \exp(-(-\ln(p_i))^\gamma) & \text{for gains} \\ \exp(-(-\ln(p_i))^\delta) & \text{for losses} \end{cases} \quad (7)$$

where γ and δ are positive parameters.

2.2 TODIM Method

TODIM is similar to CPT in that it also assumes that decision makers are more sensitive to losses than to gains and make their decisions based on certain reference points. The TODIM method involves more determinants, fewer parameters and simpler computations than CPT.

The main improvement of the TODIM method is the reformulation of the value function in CPT to include additional determinants (including qualitative ones) and reflect the loss aversion of decision makers and the dependence of choices on the reference point. The TODIM method can be summarized as follows [32, 40]:

Step 1: Determine the relevant criteria, construct a vector of normalized weights for these criteria using the Delphi method or surveys, and select the criterion with the highest weight as the reference.

Step 2: Rate each alternative for each criterion, normalize, and construct an evaluation matrix; an example is shown in Table 1:

Table 1. Evaluation matrix for the TODIM method

Alternative	Criteria				
	l_1	l_k	l_m
1	p_{11}	p_{1k}	p_{1m}
.....
n	p_{n1}	p_{nk}	p_{nm}

where n and m are the numbers of alternatives and criteria, respectively.

Step 3: Construct a value matrix using the evaluation matrix and the weights for the criteria:

$$V_{ij} = \sum_{l=1}^m \Phi_l(i, j) \forall (i, j) \tag{8}$$

$$\Phi_l(i, j) = \begin{cases} \left(\frac{w_{rl}(P_{il} - P_{jl})}{\sum_{l=1}^m w_{rl}} \right)^{\frac{1}{2}} & \text{if } (P_{il} - P_{jl}) > 0 \\ 0 & \text{if } (P_{il} - P_{jl}) = 0 \\ -\frac{1}{\mu} \left(\frac{\sum_{l=1}^m w_{rl}(P_{jl} - P_{il})}{w_{rl}} \right)^{\frac{1}{2}} & \text{if } (P_{il} - P_{jl}) < 0 \end{cases} \tag{9}$$

where V_{ij} represents the preference for alternative i over alternative j , m is the number of criteria, l denotes a specific criterion, $w_{rl} = w_l/w_r$ where r denotes the reference criterion, w_l means the weight of criterion l , p_{il} and p_{jl} are the ratings of alternatives i and j with respect to criterion l , respectively, and μ is a positive parameter.

To maintain consistency of the value functions in CPT and the TODIM method, Eq. (9) can be expressed as follows:

$$\Phi_l(i, j) = \begin{cases} \left(\frac{w_{il}(P_{il} - P_{jl})}{\sum_{l=1}^m w_{il}} \right)^\alpha & \text{if } (P_{il} - P_{jl}) > 0 \\ 0 & \text{if } (P_{il} - P_{jl}) = 0 \\ -\lambda \left(\frac{\sum_{l=1}^m w_{il}(P_{jl} - P_{il})}{w_{il}} \right)^\beta & \text{if } (P_{il} - P_{jl}) < 0 \end{cases} \quad (10)$$

where $\alpha = \beta = 0.5$ and $\lambda = 1/\mu$.

Step 4: Calculate the normalized value of each alternative using the following equation:

$$\xi_i = \frac{\sum_{j=1}^n \delta(i, j) - \min \sum_{j=1}^n \delta(i, j)}{\max \sum_{j=1}^n \delta(i, j) - \min \sum_{j=1}^n \delta(i, j)}, \quad (11)$$

where ξ_i represents the normalized value of alternative i .

A set of normalized values is calculated from Eq. (11). The TODIM method assumes that the decision maker would prefer the alternative with the highest normalized value.

3 Model and Algorithm

3.1 Model

An improved stochastic traffic assignment model with variable demands was formulated based on the TODIM method. Specifically,

$$q_w = \bar{q}_w P_w^c \quad (12)$$

$$P_w^c = \frac{\exp(\theta V_w^c)}{\sum_{i=1}^n \exp(\theta V_w^i)} \quad (13)$$

$$V_w^c = \frac{\sum_{j=1}^n V_{cj}}{n} \quad (14)$$

$$V_{cj} = \sum_{l=1}^m \Phi_l(c, j) \quad (15)$$

$$\Phi_l(c, j) = \begin{cases} \left(\frac{w_{il}(P_{cl} - P_{jl})}{\sum_{l=1}^m w_{il}} \right)^{0.5} & \text{if } (P_{cl} - P_{jl}) > 0 \\ 0 & \text{if } (P_{cl} - P_{jl}) = 0 \\ -\frac{1}{\mu} \left(\frac{\sum_{l=1}^m w_{il}(P_{jl} - P_{cl})}{w_{il}} \right)^{0.5} & \text{if } (P_{cl} - P_{jl}) < 0 \end{cases} \quad (16)$$

$$f_{wr} = q_w Pr_{wr} \quad (17)$$

$$x_a = \sum_{w \in W} \sum_{r \in R} f_{wr} \delta_{wr}^a \quad (18)$$

$$\delta_{wr}^a = \begin{cases} 1, & \text{if route } r \text{ between OD pair } w \text{ contains link } a \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

$$f_{wr} > 0 \quad (20)$$

$$q_w + \hat{q}_w = \bar{q}_w \quad (21)$$

$$q_w = \sum_{r \in R} f_{wr} \quad (22)$$

$$Pr_{wr} = \frac{\exp(\kappa V_{wr})}{\sum_{r' \in R} \exp(\kappa V_{wr'})} \quad (23)$$

$$T_w = \min_r \{ T_{wr} | P \{ t_{wr} < T_{wr} = E(t_{wr}) + \mu \sigma_{wr}^t \} \geq \rho^t \} \quad (24)$$

$$CO_w = \min_r \{ CO_{wr} | P \{ co_{wr} < CO_{wr} = E(c_{wr}) + \mu \sigma_{wr}^{co} \} \geq \rho^{co} \} \quad (25)$$

$$V_{wr} = \tau V_{wr}^t + (1 - \tau) V_{wr}^{co} \quad (26)$$

$$v_{wr}(t_{wi}) = \begin{cases} (T_{bw} - t_{wi})^\alpha & \text{if } T_{bw} - t_{wi} \geq 0 \\ -\lambda(-(T_{bw} - t_{wi}))^\beta & \text{if } T_{bw} - t_{wi} < 0 \end{cases} \quad (27)$$

$$v_{wr}(co_{wi}) = \begin{cases} (CO_{bw} - co_{wi})^\alpha & \text{if } CO_{bw} - co_{wi} \geq 0 \\ -\lambda(-(CO_{bw} - co_{wi}))^\beta & \text{if } CO_{bw} - co_{wi} < 0 \end{cases} \quad (28)$$

$$V_{wr}^t = \int_{T_{bw}}^{+\infty} -\frac{dw(1-F(x))}{dx} v(x) dx + \int_{-\infty}^{T_{bw}} \frac{dw(F(x))}{dx} v(x) dx \quad (29)$$

$$V_{wr}^{co} = \int_{CO_{bw}}^{+\infty} -\frac{dw(1-F(x))}{dx} v(x) dx + \int_{-\infty}^{CO_{bw}} \frac{dw(F(x))}{dx} v(x) dx \quad (30)$$

where the variables are defined as follows:

- \bar{q}_w – total demand (i.e., number of trips) between origin-destination (OD) pair w ,
- \hat{q}_w – demand for public transit (i.e., the subway) between OD pair w ,
- q_w – demand for driving privately owned vehicles between OD pair w ,
- p_w^c – probability of choosing to drive privately owned vehicle between OD pair w ,
- θ – scale for the value of a travel mode choice,
- V_w^c – value of driving between OD pair w ,
- f_{wr} – traffic volume of privately owned vehicles using route r between OD pair w ,
- Pr_{wr} – probability of choosing to drive a privately owned vehicle on route r between OD pair w ,
- κ – scale for the prospect of a route choice,

- x_a – traffic volume on link a ,
 T_w – reference point for the travel time between OD pair w ,
 CO_w – reference point for the travel cost between OD pair w ,
 V_{wr} – general prospect of route r between OD pair w ,
 V_{wr}^t – travel time prospect of route r between OD pair w ,
 V_{wr}^{co} – travel cost prospect of route r between OD pair w ,
 τ – weight on travel time for a route choice.

Equations (12) and (13) define the modal split model and can also be regarded as variable demand functions. The probability is expressed in logit form. Equations (14)–(16) are used to calculate the utilities of the travel modes and are based on the TODIM. Equations (17)–(23) are the stochastic user equilibrium model, of which Eqs. (22) and (23) are the logit-based loading model with the general prospect of each route. Equations (24) and (25) define the reference point constraint conditions with respect to the travel time and the travel cost, respectively. Equation (26) generates the general prospect. Equations (27) and (28) are the value functions with respect to the travel time and the travel cost, respectively, and Eqs. (29) and (30) give the prospects of the travel time and the travel cost, respectively.

The BPR function is assumed for the impedance function:

$$t_a = t_a^0 \left[1 + 0.15 * \left(\frac{x_a}{cap_a} \right)^4 \right] \quad (31)$$

where t_a is the link a ' travel time, t_a^0 is this link' free-flow travel time, x_a is its traffic flow, and cap_a is its capacity.

3.2 Algorithm

A heuristic algorithm based on the method of successive averages (MSA) was used to find the traffic assignments and is described in this section. Because this algorithm is based on MSA, the step size is predetermined.

4 Numerical Example

4.1 Criteria Determination

The first step is to establish the criteria that influence travel mode choices. In this example, the criteria can be divided into three main categories: the characteristics of the travelers, the characteristics of trips, and the characteristics of each travel mode. For convenience, only the characteristics of the travel modes are considered in the numerical example. It can be observed that the travel time and the travel cost are quantitative variables, while the others are qualitative. For the qualitative variables, a scoring method was employed to evaluate the various travel modes.

4.2 Example Network

The Nguyen and Dupuis network was used to test the validity of the improved stochastic equilibrium model with variable demands. This network consists of 13 nodes, 19 links and 4 origin-destination (OD) pairs. Furthermore, we assumed that each origin and destination were connected by a subway system. The road network and the subway network are independent because they have different spatial configurations. The example network is shown in Fig. 1.

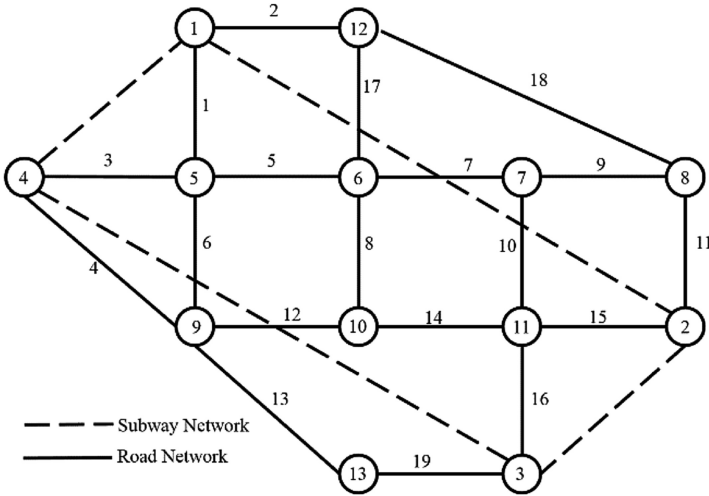


Fig. 1. Example travel network.

There are 4 OD pairs, (1, 2), (1, 3), (4, 2) and (4, 3), in the network. The total demand levels for the OD pairs were $\bar{q}_{12} = 800$, $\bar{q}_{13} = 1000$, $\bar{q}_{42} = 800$ and $\bar{q}_{43} = 1000$. The travel time of the subway network are 40 min (OD:1,2), 55 min (OD:1,3), 55 min (OD:4,2) and 70 min (OD:4,3) respectively. Because subway systems operate on a fixed schedule, the travel time by subway for each OD pair was fixed. Furthermore, a fixed ticket price, 2 RMB, was assumed.

The characteristics of the routes between each OD pair are shown in Table 2, where σ^2 is the variability parameter, i.e., the variance per unit of the travel time or the travel cost for each route. In this example, it was assumed that the travel time and the travel cost had the same variability parameters.

4.3 Criterion Weights and Evaluation Matrix

To obtain the weights on the criteria and the evaluation matrix, a survey was conducted through the Internet. In total 112 questionnaires were collected during which 111 were completed and have no obvious mistake. So these 111 samples are considered as valid. All respondents were well educated, and the majority were of ages 18 to 30 years (88%). Approximately 60% of the respondents were male, and 61% possessed a

Table 2. Characteristics of the routes between OD pairs

OD	Route	Link ID	σ^2	OD	Route	Link	σ^2
(1, 2)	1	2,18,11	0.5	(1, 3)	1	2,17,7,10,16	0.3
	2	2,17,7,9,11	0.01		2	2,17,8,14,16	0.001
	3	2,17,7,10,15	0.15		3	1,5,7,10,16	0.3
	4	2,17,8,14,15	0.2		4	1,5,8,14,16	0.5
	5	1,5,7,9,11	0.1		5	1,6,12,14,16	0.02
	6	1,5,7,10,15	0.3		6	1,6,13,19	0.2
	7	1,5,8,14,15	0.001				
	8	1,6,12,14,15	0.01				
(4, 2)	1	4,12,14,15	0.1	(4, 3)	1	4,13,19	0.6
	2	3,5,7,9,11	0.4		2	4,12,14,16	0.1
	3	3,5,7,10,15	0.001		3	3,6,13,19	0.3
	4	3,5,8,14,15	0.35		4	3,5,7,10,16	0.05
	5	3,6,12,14,15	0.02		5	3,5,8,14,16	0.01
					6	3,6,12,14,16	0.1

driver's license. From the results, it can be concluded that most of the respondents were college students and young, white-collar workers who had very positive opinions of traveling both by the subway system and by automobile. Table 3 shows the values of the criterion weights obtained from the averages of the survey results. Each respondent was asked to grade the criteria on a scale from 1 (least important) to 5 (most important). The results showed that punctuality was the most important criterion for commuting, followed by travel time, safety, comfort and travel cost, in that order.

Table 3. Criterion weights.

Criterion	Travel time	Travel cost	Safety	Comfort	Punctuality
Average score	3.64	3.15	3.34	3.25	4.12
Normalization	0.21	0.18	0.19	0.19	0.24

Table 4 gives the results for the qualitative criteria related to driving privately owned vehicles and riding the subway system. The results showed that most respondents rated privately owned vehicles highest on comfort and lowest on punctuality, whereas the ratings for the subway system were the opposite.

Table 4. Privately owned vehicle and subway characteristics (Normalization).

Criteria	Safety	Comfort	Punctuality
Privately owned vehicle	3.51 (0.46)	4.28 (0.59)	3.41 (0.44)
Subway	4.10 (0.54)	3.03 (0.41)	4.29 (0.56)

4.4 Parameters

The model involves many parameters, and it is necessary to explain the significance of these parameters. Regarding the CPT parameters ($\lambda, \alpha, \beta, \gamma$ and δ), the values obtained in previous research [39] were used in this example. The ratio $1/\mu$ in the TODIM method has the same definition as λ in CPT, as shown in Eqs. (9) and (10), and thus we set $1/\mu = \lambda$. The parameters ρ^t and ρ^c represent travelers' attitudes toward risk with regard to travel time and travel cost, respectively. For convenience, it was assumed that $\rho^t = \rho^c = \rho$. The parameter τ is the weight on travel time for a route choice, which was determined from the results of the survey.

4.5 Results

According to Kahneman and Tversky, decisions are reference dependent. Figure 2 shows the relationship between the demand for privately owned vehicle travel and the parameter ρ . Figure 3 takes OD pair (1, 2) as an example to show the relationship between the prospect for each route and the parameter ρ .

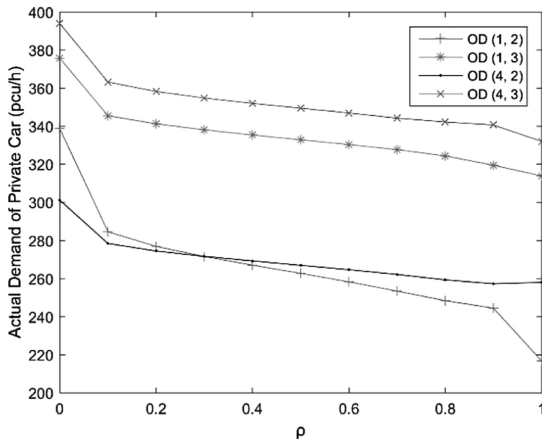


Fig. 2. Demand for driving vs. the reference point.

In Fig. 2, the demand for privately owned vehicle travel decreases as the value of ρ increases. The reason for this trend is that as the value of ρ increases, the travelers' reference points gradually increase, which increases the travel time and cost of driving (privately owned vehicles), and thus its attractiveness decreases. The parameter ρ reflects travelers' risk aversion regarding travel time or travel cost, and a larger value of ρ indicates a more risk-averse attitude. It can be concluded that more risk-averse travelers tend to choose to ride the subway rather than drive. In Fig. 3, the prospect of each route increases as the value of ρ increases. This trend occurs mainly because an increase in the travelers' reference points leads to a shift from routes with losses to routes with gains, causing the prospects of these routes to increase. The variations in the route prospects shown in Fig. 3 reveal a similar trend, however the differences

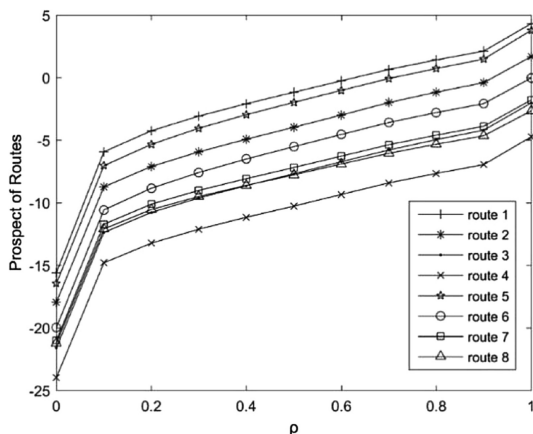


Fig. 3. Route prospects for OD pair (1, 2) vs. ρ .

among routes' prospect do vary based on ρ , especially when ρ is shift from below 0.1 to above 0.1. Because only the differences are significant in a discrete choice model, a conclusion can be conducted that the travelers' route choices are reference dependent.

Figure 4 shows the relationship between the demand for driving and the parameter θ . The relationship between the route choice probability and the parameter κ for OD pair (1, 2) is shown in Fig. 5.

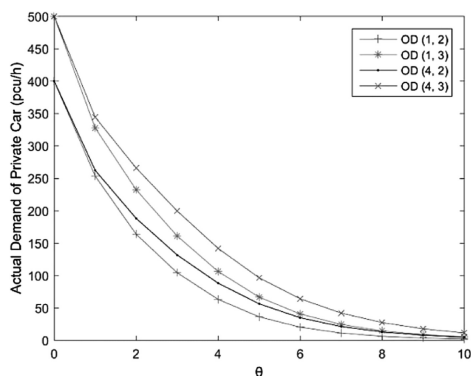


Fig. 4. Demand for driving vs. θ .

The parameter θ is the scale for the travel modes, reflecting travelers' preferences for the various travel modes. Figure 4 indicates that the demand for driving (i.e., travel by privately owned vehicles) decreases as the value of θ increases. The demand for driving and riding the subway are equal when θ is equal to 0, i.e., travelers have no preference for either mode. As the value of θ approaches positive infinity, travelers have a strong preference and always choose to use the subway rather than to drive.

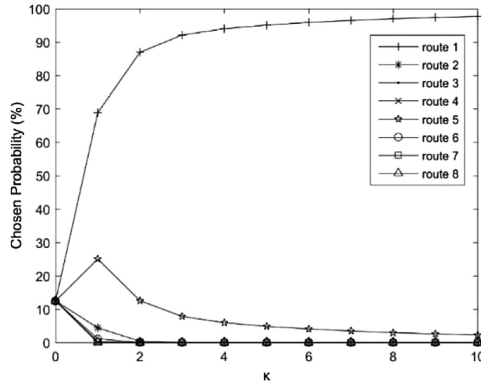


Fig. 5. Route choice probability for OD pair (1, 2) vs. κ .

Similarly, the parameter κ is the scale of travel route prospects, reflecting travelers’ preferences for the various travel routes. As Fig. 5 shows, travelers have no route preference when the value of κ is equal to 0. As the value of κ approaches positive infinity, however, travelers have a strong preference for the route with the highest prospect.

The preceding analysis shows that the parameters θ and κ are critical to the model. Therefore, the values of these parameters must be estimated carefully based on data obtained from decision makers through surveys. In this example, these two parameters can be roughly valued in (0, 6) and (0, 1), respectively.

5 Conclusion

The traditional traffic assignment model is based on the principle of random utility maximization which assumes that decision makers prefer the alternative with the highest utility. This study modeled mode and route choices behavior considering travelers’ bounded rationality. An improved stochastic traffic model with multi-modes was formulated, and a heuristic algorithm was developed. A numerical example was presented to confirm the validity of this model. Analysis of this numerical example revealed that the reference dependence proposed by Kahneman and Tversky does exist in travelers’ choices of various routes and the scaling parameters in the model play an important roles, which should be estimated using real data.

Several unresolved problems remain. First, although the validity of the model was confirmed, this model has not been applied to a complex road network. An efficient algorithm for finding routes in a complex road network is required before the model can be applied to such cases. Second, this study mainly focused on privately owned vehicles and not on public transit. Future studies could address the multi-modal traffic assignment problem.

Acknowledgments. This research was supported by the National Natural Science Foundation of China (Number 71861032), National Natural Science Foundation of Xinjiang (Number 2018D01C071) and the doctoral research fund of Xinjiang University.

Ethical Approval. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent. Informed consent was obtained from all individual participants included in the study.

References

1. Ben-Akiva, M.E., Lerman, S.R.: *Discrete Choice Analysis: Theory and Application to Travel Demand*, vol. 9. MIT Press, Cambridge (1985)
2. Yai, T., Iwakura, S., Morichi, S.: Multinomial probit with structured covariance for route choice behavior. *Transp. Res. B Methodol.* **31**, 195–207 (1997)
3. Vovsha, P., Bekhor, S.: Link-nested logit model of route choice: overcoming route overlapping problem. *Transp. Res. Rec.* **1645**, 133–142 (1998)
4. Bhat, C.R.: Analysis of travel mode and departure time choice for urban shopping trips. *Transp. Res B Methodol.* **32**, 361–371 (1998)
5. Simon, H.A.: A behavioral model of rational choice. *Q. J. Econ.* **69**, 99–118 (1955)
6. Nakayama, S., Kitamura, R., Fujii, S.: Drivers' learning and network behavior: dynamic analysis of the driver-network system as a complex system. *Transp. Res. Rec.* **1676**, 30–36 (1999)
7. Nakayama, S., Kitamura, R., Fujii, S.: Drivers' route choice rules and network behavior: do drivers become rational and homogeneous through learning? *Transp. Res. Rec.* **1752**, 62–68 (2001)
8. Gifford, J.L., Checherita, C.: Bounded rationality and transportation behavior: lessons for public policy. In: *TRB 86th Annual Meeting Compendium of Papers CD-Rom (No. 07-2451)*, Washington, DC (2007)
9. Allais, M.: Le comportement de L'Homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica J. Econometric Soc.* **20**, 503–546 (1953)
10. Ellsberg, D.: Risk, ambiguity, and the savage axioms. *Q. J. Econ.* **75**, 643–669 (1961)
11. Simon, H.A.: *Bounded Rationality. Utility and Probability*. Palgrave Macmillan, London (1987)
12. Gigerenzer, G., Goldstein, D.G.: Reasoning the fast and frugal way: models of bounded rationality. *Psychol. Rev.* **103**, 650–669 (1996)
13. Conlisk, J.: Why bounded rationality? *J. Econ. Lit.* **34**, 669–700 (1996)
14. Kahneman, D.: A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.* **58**, 697–720 (2003)
15. Tang, T., Huang, H., Shang, H.: Influences of the driver's bounded rationality on micro driving behavior, fuel consumption and emissions. *Transp. Res. D* **41**, 423–432 (2015)
16. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. *Econometrica* **47**, 263–291 (1979)
17. Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertainty* **5**, 297–323 (1992)

18. Rieger, M.O., Wang, M.: Cumulative prospect theory and the St. Petersburg paradox. *Econ. Theory* **28**, 665–679 (2006)
19. Dhami, S., Al-Nowaihi, A.: Why do people pay taxes? Prospect theory versus expected utility theory. *J. Econ. Behav. Organ.* **64**, 171–192 (2007)
20. Pan, X.F., Zuo, Z.: A stochastic user equilibrium model and optimal congestion pricing with prospect theory. *Procedia Soc. Behav. Sci.* **138**, 127–136 (2014)
21. Senbil, M., Kitamura, R.: Reference points in commuter departure time choice: a prospect theoretic test of alternative decision frames. *J. Intell. Transp. Syst.* **8**, 19–31 (2004)
22. Jou, R.C., Kitamura, R., Weng, M.C., Chen, C.C.: Dynamic commuter departure time choice under uncertainty. *Transp. Res. A Policy Pract.* **42**, 774–783 (2008)
23. Liu, D., Lam, W.H.: Modeling the effects of population density on prospect theory-based travel mode-choice equilibrium. *J. Intell. Transp. Syst.* **18**, 379–392 (2014)
24. Connors, R.D., Sumalee, A.: A network equilibrium model with travellers' perception of stochastic travel times. *Transp. Res. B* **43**, 614–624 (2009)
25. Xu, H., Lou, Y., Yin, Y., Zhou, J.: A prospect-based user equilibrium model with endogenous reference points and its application in congestion pricing. *Transp. Res. B Methodol.* **45**, 311–328 (2011)
26. Avineri, E., Bovy, P.: Identification of parameters for a prospect theory model for travel choice analysis. *Transp. Res. Rec.* **2082**, 141–147 (2008)
27. Jou, R.C., Chen, K.H.: An application of cumulative prospect theory to freeway drivers' route choice behaviours. *Transp. Res. A Policy Pract.* **49**, 123–131 (2013)
28. Nwogugu, M.: Towards multi-factor models of decision making and risk: a critique of prospect theory and related approaches, Part I. *J. Risk Finance* **6**, 150–162 (2005)
29. Nwogugu, M.: Towards multi-factor models of decision making and risk: A critique of prospect theory and related approaches, Part II. *J. Risk Finance* **6**, 163–173 (2005)
30. Nwogugu, M.: Towards multi-factor models of decision making and risk: a critique of prospect theory and related approaches, Part III. *J. Risk Finance* **6**, 267–274 (2005)
31. Timmermans, H.: On the (ir)relevance of prospect theory in modelling uncertainty in travel decisions. *EJTIR* **4**, 368–384 (2010)
32. Gomes, L., Lima, M.: TODIM: basics and application to multicriteria ranking of projects with environmental impacts. *Found. Comput. Decis. Sci.* **16**, 113–127 (1992)
33. Gomes, L., Lima, M., Maranhão, F.: Multicriteria analysis of natural gas destination in Brazil: an application of the TODIM method. *Math. Comput. Modell.* **50**, 92–100 (2009)
34. Tseng, M.L., Lin, Y.H., Tan, K., Chen, R.H., Chen, Y.H.: Using TODIM to evaluate green supply chain practices under uncertainty. *Appl. Math. Modell.* **38**, 2983–2995 (2014)
35. Giulio, E.C., Armando, C., Stefanode, L.: Stochastic equilibrium assignment with variable demand: theoretical and implementation issues. *Eur. J. Oper. Res.* **241**, 330–347 (2015)
36. Tversky, A., Kahneman, D.: Loss aversion in riskless choice: a reference-dependent model. *Q. J. Econ.* **106**, 1039–1061 (1991)
37. Prelec, D.: The probability weighting function. *Econometrica* **66**, 497–527 (1998)
38. Ingersoll, J.: Non-monotonicity of the Tversky-Kahneman probability-weighting function: a cautionary Note. *Eur. Finan. Manag.* **14**, 385–390 (2008)
39. Zuo, Z., Pan, X.F., Liu, K.: Parameter calibration in cumulative prospect theory for travelers' route choice behavior. In: 15th COTA International Conference of Transportation Professionals (CICTP 2015), pp. 2696–2708. American Society of Civil Engineers, Reston (2015)
40. Gomes, L.: An application of the TODIM method to the multicriteria rental evaluation of residential properties. *Eur. J. Oper. Res.* **193**, 204–211 (2009)



A Map Matching Algorithm Combining Twice Gridding and Weighting Factors Methods

Ketu Cao, Lixiao Wang^(✉), Zhi Zuo, and Xiaohui Sun

School of Civil Engineering and Architecture, Xinjiang University,
Urumqi, Xinjiang, China
xjwanglx@foxmail.com

Abstract. Current map matching algorithms suffer the problem that matching accuracy and matching efficiency cannot be achieved both in the face of massive floating car data. A novel map matching algorithm combining twice gridding and weighting factors methods has been proposed in this study, selection of candidate road segments and determination of the shortest path are based on twice gridding method; Factors of driving direction and trajectory angle are served as improvement of weighting factors method, and matching accuracy rate in specific situations such as parallel road segments, intersection areas, intensive road segments and large positioning errors has been improved. Meanwhile, to avoid driving direction and trajectory angle failing at low speeds, which results in interference problems with map matching, the weights of driving directions and trajectory angle are dynamically adjusted by instantaneous speed and interval length between the trajectories. Through map matching case study of actual data, results show that the improved weighting factors method in this study performs outstandingly in improving the matching accuracy rate, and combining the method of twice gridding effectively improves matching efficiency in principle, which makes the map matching algorithm in this study can balance the matching accuracy and matching efficiency well when facing massive floating car data.

Keywords: Map matching · Weighting factors · Twice gridding · Floating car data

1 Introduction

Nowadays, urban congestion has become an increasingly severe problem, and have caused serious impacts on urban economic development. The grasp of urban traffic conditions can provide reliable technical support for the strategy formulation to alleviate traffic congestion. As a main way of traffic information collection, floating car technology can provide data support for urban traffic analysis. Map matching is the basis of data processing for floating car data, and its accuracy has a significant impact on the results of data analysis.

The location information is acquired by the position sensing device based on mobile phone, on board unit (OBU) etc., which is the original GPS positioning data, coupled with road network data, constraints are established based on correlation

between GPS positioning data and network data, then a travel path that conforms to the road network structure can be found. This technology is called map matching [1]. The constraint function system formed by the integration of multiple constraint functions that are established to reduce data errors is called a map matching algorithm. According to the basic principles, map matching algorithm can be divided into three types: simple map matching algorithm, weighting factors map matching algorithm and advanced map matching algorithm [1].

The simple map matching algorithm is dependent on a single element with fewer constraints to determine matching results. The original simple map matching algorithm is a point-to-point map matching algorithm proposed by Kim [2] and Bernstein [3]. The principle is to select the network node which is closest to GPS positioning point as the matching node. White [4], Toylar [5] and Miwa [6] improved the simple map matching algorithm, proposing point-to-line, line-to-line, map matching algorithms based grid. Because the simple map matching algorithm can not fully achieve high matching accuracy, to solve this problem, Greenfeld [7] proposed an weighting factors map matching algorithm considering driving direction, distance, angle (the angle between the line direction from GPS positioning point to intersection and the direction of road segment). Quddus [8] improved Greenfeld [7]'s method, making the algorithm more simpler and applicable. Although the weighting factors algorithm can achieve higher precision, there are still a certain number of matching errors in parallel road segments, intersections, intensive road segments network and large error of positioning data. In order to achieve higher matching accuracy, many scholars have proposed the advanced map matching algorithms that are different from the basic principles of the first two map matching algorithm. For example, Syed [9] and Ren [10] proposed a map matching algorithm based on fuzzy logic; Ren [11] and Wu [12] proposed a map matching algorithm based on Markov model.

The performance of map matching algorithms requires high matching accuracy and high matching efficiency. The simple map matching algorithm can't meet the matching precision requirement. However, most weighting factors map matching algorithm applies the probability ellipse or quadrilateral error region to select the candidate path, for each GPS positioning point, the candidate range needs to be recalculated, it reduces matching efficiency, and they could not meet the accuracy requirement on specific road network due to the static weight of the algorithm. The advanced map matching algorithm often need the calculation of multiple GPS positioning points to determine one GPS point actual location, which means they are more complicated in read data and calculation. Those complexity not only reduce advanced algorithm efficiency but also make algorithm's applicability has limitations.

In order to solve the fore-mentioned problems, this study proposes a novel map matching algorithm that combines twice gridding and improved weighting factors method. The algorithm acquires candidate road segments based on small grid, it only needs to calculate the candidate road segment selection range once; Weighting factors map matching algorithm is improved by dynamically adjusting the factors weights by instantaneous speed, interval length, etc., those problems of parallel road segment jump, intersection mismatch, intensive road segment mismatch and poor fault tolerance when the positioning error is large have been overcome, and the optimal matching road segment can be determined; Large grid is used to match the shortest path.

Following section introduces the principle of matching algorithm and the improvement of the factor weighting method in this study, then the next is the case study, matching example verification and result analysis is shown, and final section is conclusion. The algorithm proposed in this study has good performance in both matching accuracy and efficiency, and has great value on practical application.

2 Principle of Matching Algorithm

The main contents of the map matching algorithm combining twice gridding and weighting factors methods proposed in this study are: 1. Obtaining candidate segments based on small grids; 2. Based on improved factors weighting algorithm, local map matching determines matching road segments; 3. The large grid determines the candidate path range, and the shortest path is selected as the matching path. Detailed flow chart of the algorithm is given in Fig. 1.

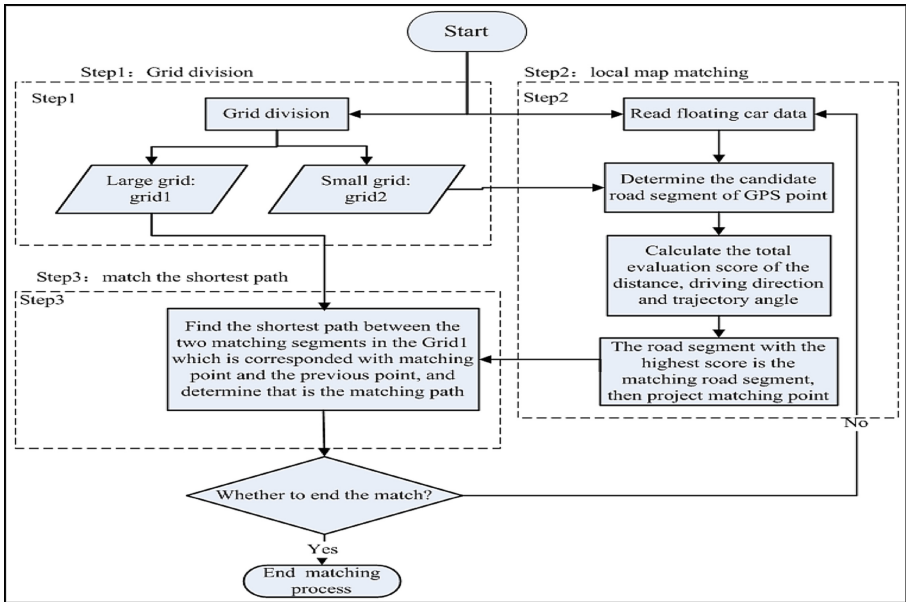


Fig. 1. Map matching algorithm flow chart

2.1 Grid Division

Conventional weighting factors map matching algorithm [2–8] uses probabilistic ellipse or quadrilateral to acquire candidate segments, which reduces computational efficiency. Cao [13], Zhao [14] proposed a gridding map method, using grid as error region to determine candidate road segments, which improving the algorithm efficiency. Cao’s twice gridding method [13] is applied in grid division to determine the

candidate road segments in this study. The twice gridding method refers to dividing the research area into $M \times N$ equal-sized grids, the attributes of the intersected road segments can be found in the grid, searching capability is strengthened. There are two types grids: large grid (grid1) and small grid (grid2). Grid1 determines candidate path range; Grid2 determines candidate road segment range in the algorithm.

2.2 Local Map Matching

The process of local map matching consists of two main parts: finding candidate road segments and obtaining matching road segments.

Finding candidate road segments: Finding the Grid2 which containing positioning point, obtaining road segments from this Grid2 and its eight neighboring Grid2, selecting candidate road segments based on the distance from the positioning point to the road segment and the length of the road segment, forming a candidate road segment set [13].

This study uses distance, driving direction and trajectory angle as the factors of weighting factors algorithm. On this basis, instantaneous speed attribute, which has greater correlation with the driving direction, is applied to control driving direction factor, and driving distance attribute, which has greater correlation with the trajectory angle factor, is applied to control the weight of trajectory angle factor, then matching precision can be improved in some conditions (parallel road sections, intensive road sections and large positioning errors, etc.), which improves the precision and efficiency of algorithm to some extent. Detailed formula of improved weighting factor algorithm proposed in this study is shown as follows:

$$w_{sum} = \alpha_{distance}w_{distance} + \alpha_{direction}w_{direction} + \alpha_{angle}w_{angle} \quad (1)$$

w_{sum} : Total similarity score of distance factor, driving direction factor and trajectory angle factor.

$w_{distance}, w_{direction}, w_{angle}$: Similarity scores for distance, driving direction, and trajectory angle factor respectively.

$\alpha_{distance}, \alpha_{direction}, \alpha_{angle}$: Value of weighting factors for distance, driving direction, and trajectory angle factor respectively.

The trajectory angle factor and the driving direction factor are easily affected by the low driving speed in offline map matching. In contrast, the distance factor is more stable than the trajectory angle factor and the driving direction factor in such cases, so the $\alpha_{distance}$ is 1.2. The specific calculation formula of $w_{distance}$ is shown as follows:

$$w_{distance} = \frac{1}{\left(1 + \left(\frac{d}{e}\right)^2\right)} \quad (2)$$

d : Distance between the positioning point and candidate road segment.

e : Radius of floating car data error.

It is proposed that the driving direction accuracy in the collected information has apparent relation with the floating vehicle instantaneous speed in existing researches [15]. With higher instantaneous speed value, the collected driving direction is closer to actual driving direction. Consequently, this study uses the instantaneous speed factor to adjust the calculated value of $\alpha_{direction}$ to improve the calculation of driving direction factor. The specific calculation formula is shown as follows:

$$\alpha_{direction} = \begin{cases} 1 & v_s > v_2 \\ 1 \times \left(\frac{v_2 - v_s}{v_2 - v_1} \right) & v_2 > v_s > v_1 \\ 0 & v_1 > v_s \end{cases} \quad (3)$$

$$w_{direction} = \cos\left(\frac{\Delta\alpha}{2}\right) \quad (4)$$

v_2 : Preset speed threshold value, when the instantaneous speed value is higher than the threshold, the driving direction factor has higher reliability.

v_1 : Preset speed threshold value, when the instantaneous speed value is less than the threshold, the driving direction factor has lower reliability.

v_s : Instantaneous speed value of floating car.

$\Delta\alpha$: The angle between driving direction of floating car and direction of road segment.

The trajectory angle factor are easily affected by elements such as collection frequency, driving speed and interval length of the adjacent floating car positioning points. It can also be more credible with reduction of collection frequency within certain range, increase of driving speed and extension of interval length [16]. Therefore selecting interval length to adjust the value of α_{angle} improves the calculation of α_{angle} , it can avoid the effect of trajectory angle factor in local map matching process when the trajectory angle feature is at low precision. The trajectory angle factor is calculated as follows:

$$\alpha_{angle} = \begin{cases} 1 & s \geq s_0 \\ \sqrt{\frac{s}{s_0}} & s_0 > s \end{cases} \quad (5)$$

$$w_{angle} = \cos\left(\frac{\Delta\beta}{2}\right) \quad (6)$$

s : Interval length of time-adjacent positioning points.

s_0 : Present maximum interval length of time-adjacent positioning points.

$\Delta\beta$: The angle between the direction of road segment and the direction from last matched positioning point to pre-match positioning point.

According to the above formula, for a single positioning point, the total similarity score for each candidate road segment can be obtained, and the road segment with the highest total similarity score is selected as the matching road segment.

2.3 Match the Shortest Path

The algorithm in this study utilizes method of matching shortest path part in algorithm [13], in which Grid1 delineated by the twice gridding is treated as candidate path searching region, then the Grid1 corresponding to the adjacent GPS points will be combined as the path searching range, and the Dijkstra algorithm is applied to find the shortest path to determine the matching path. Based on above method, additional method of dividing the adjacent GPS points whose driving distance interval is too long into different trajectory chains is applied, as a improved measure to avoid the unnecessary matching error and computation time caused by long driving distance.

3 Case Study

3.1 Introduction of Data

The data used in this study is selected from floating car data of Urumqi, Xinjiang province, and data collection time is January 2nd, 2016, with collection cycle of 30 s. Basic information of floating car data is shown in Table 1.

Table 1. Basic information of floating car data

Collecting date	Collecting time	License Plate	Longitude	Latitude	Speed (km/h)	Driving direction
20160102	215422	新AN542*	43.8235764	87.6345465	29	186
20160102	215452	新AN542*	43.8216593	87.6342880	24	183

Note: The first two digits of collecting time indicate the hour in 24-h, the next two digits of collecting time indicate minute, and the last two digits of collecting time indicate second.

3.2 Parameter Setting

In this paper, the urban area is divided into 140×160 Grid1 with a geometrical parameter of $285 \text{ m} \times 250 \text{ m}$. When searching for the shortest path, it will expand to the larger grid with geometrical parameter of $855 \text{ m} \times 750 \text{ m}$. Based on the grid structure, it can be seen that at least the shortest path can be found when the adjacent GPS points are within 500 m (the speed value is below 60 km/h). According to the speed distribution (Table 2), it is known that at least 96.44% of the shortest paths between adjacent GPS points can be found, which ensures the connectivity of the road network and the continuity of the trajectory. Urban area of Urumqi is divided into 1400×1600 Grid2 with a geometric parameters of $28.5 \text{ m} \times 25 \text{ m}$. When searching

for matching road segments, it will expand to $85.5 \text{ m} \times 75 \text{ m}$. In this paper, the floating car positioning error radius is 30 m, whereas candidate range is $85.5 \text{ m} \times 75 \text{ m}$, larger than the error range, thus the possibility of missing the correct road segment is avoided when selecting the candidate road segment.

Table 2. Instantaneous speed distribution

Instantaneous speed interval	Cumulative number	Number in intervals	Proportion in intervals	The cumulative proportion
≤ 10 (km/h)	69306	69306	39.37%	39.37%
$>10 \ \& \ \leq 20$ (km/h)	89653	20347	11.56%	50.93%
$>20 \ \& \ \leq 30$ (km/h)	112003	22350	12.70%	63.63%
$>30 \ \& \ \leq 40$ (km/h)	141662	29659	16.85%	80.48%
$>40 \ \& \ \leq 50$ (km/h)	160190	18528	10.53%	91.01%
$>50 \ \& \ \leq 60$ (km/h)	169756	9566	5.43%	96.44%
$>60 \ \& \ \leq 70$ (km/h)	174230	4474	2.54%	98.98%
>70 (km/h)	176020	1790	1.02%	100.00%

According to the collected data attribute, the floating car positioning error radius ε is 30 m. Ochieng et al. [15] proposed that when instantaneous speed value is equal to or greater than 10.8 km/h (3 m/s), it can be ensured that the difference between the driving direction of the floating car record and the actual direction is less than 30° . According to the relationship description between the actual speed value and the driving direction error value in the Ochieng’s study [15], the relationship between that two are shown in Fig. 2. Based on the variation trend, when instantaneous speed reaches 80 km/h (22.2 m/s) or higher, driving direction error is less than 5° , and direction accuracy is approximately 95%. Therefore, the speed threshold v_1 is 10.8 km/h (3 m/s), and the speed threshold v_2 is 80 km/h (22.2 m/s).

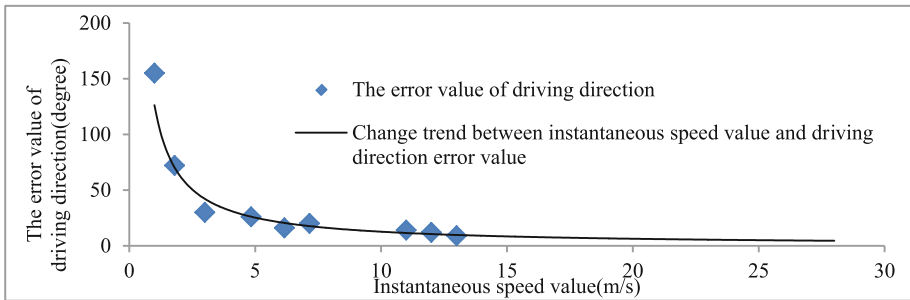


Fig. 2. Change trend between instantaneous speed value and driving direction error value

The trajectory angle factor is mainly affected by driving distance and positioning error, which is shown in Fig. 3 and Table 3. Assuming that the floating car drives in a straight road segment with an positioning error radius of 30 m. When the interval between adjacent positioning points increases, the maximum angular deviation decreases continuously until the interval length reaches 350 m, and the minimum accuracy rate reaches over 90%. However, in practice, there are driving behaviors different from going straight such as turning, then the minimum error will be less than 90%. Since the time interval of the collection is $T = 30$ s, considering the trajectory angle factor change interval should contain at least 95% of driving trajectory, therefore, interval length threshold is taken as 600 m (the speed is 70 km/h) to decrease the proportion of the trajectory angle factor in the overall evaluation. And Fan [16] compares and analyzes the relationship between the trajectory angle and the interval length, and confirms that the interval length threshold is 600 m to make the trajectory angle factor reliability.

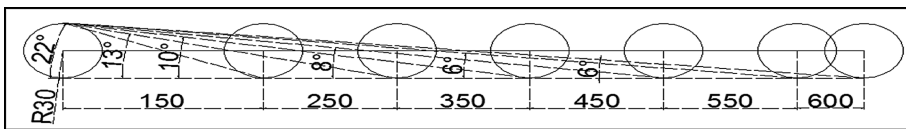


Fig. 3. The change schematic diagram of the angle and interval length on the straight road segment

Table 3. Maximum angular deviation, interval length and minimum accuracy rate change on straight road segments

The interval length	Maximum angular deviation	Minimum accuracy
150 m	22°	76%
250 m	13°	86%
350 m	8°	91%
550 m	7°	92%
600 m	6°	93%

Note: Minimum accuracy = maximum angular deviation/90°.

4 Analysis of Results

In the performance test of the map matching algorithm, it should be evaluated the matching accuracy rate.

At present, there is no uniform standard for map matching accuracy rate measurement methods, but the road segment matching accuracy rate is most measurement. The road segment matching accuracy rate refers to the ratio of correct matching number of road segments to the total matching number of road segments. Therefore, the road segment matching accuracy rate is selected as the evaluation criterion of the matching accuracy rate in this study. The data used in this paper are historical data, the actual

trajectory is unavailable, but through the timing sequence of positioning point and road network connectivity, the path contains the most positioning points and with the shortest length can be known. This paper regards this kind of path as the actual driving trajectory.

Analyzing about 1000 positioning point matching results, this algorithm's correct rate reaches 99%. The accuracy rate of normal weighting factor algorithm [8, 9, 11] is about 90%, the accuracy rate of advanced map matching algorithm [14] is about 99%. The matching accuracy rate results show that the map matching algorithm in this paper is significantly higher than the normal weighting factor algorithm, and it is similar with advanced map matching algorithm, which means that the map matching algorithm in this paper has a very high matching accuracy.

The convention weighting factor algorithm leads to the imbalance between their factors due to the static weight. In specific case such as the road network structure composed of surface and elevated road, the roads are linearly similar and parallel, and intensive road segments or large positioning error, the wrong results of road segment selection often appear, which reduces the matching accuracy. The improved algorithm can avoid those problems. In the road network shown in Fig. 4, which composed of surface and elevated road, and these roads are parallel. The middle road is an elevated road, and it is impossible to turn around. But the actual trajectory has a turning round behavior. The improved algorithm finally matched surface road segments that allows U-turn as resulting road segments and without jump selection of road segment. The matching result of algorithm in this paper is shown in Fig. 5. The circled two points in Fig. 5 (left one) is far from actually driving trajectory. But final matching road segment isn't the one belong to the right-turning road, which is the closest to the point, but belongs to the straight road segment which is actually driving trajectory. The result of Fig. 5 (right one) also proves that improved algorithm has good matching performance in intensive road segments.

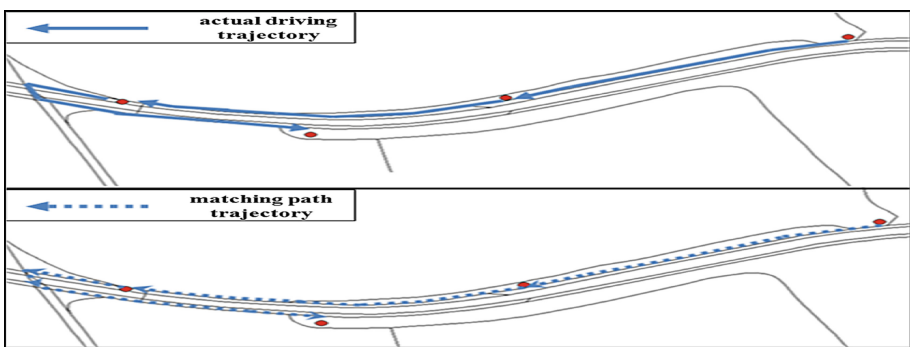


Fig. 4. Match path instance

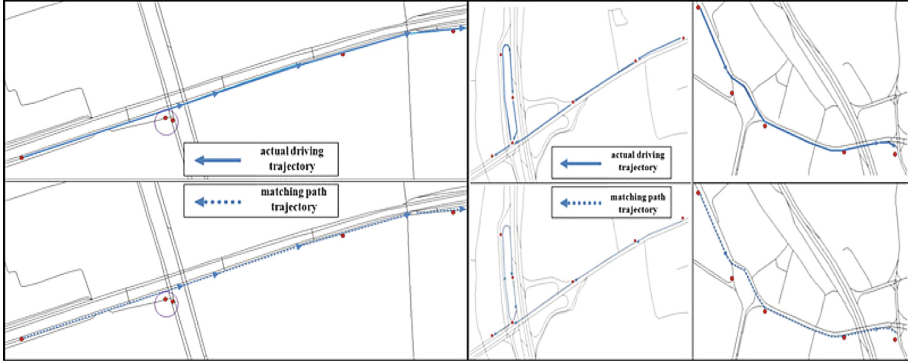


Fig. 5. Match path instance

From the matching results analysis, the improved algorithm performs excellent in the road segment matching accuracy rate on most general road network structures and the specific urban road network, even in the case of large local positioning errors. It is proved that the proposed algorithm has high overall accuracy rate.

5 Conclusion

The existing gridding map matching algorithms mostly make the shortest path as the core and cannot meet accuracy requirements. The normal weighting factors algorithm needs to construct the error region multiple times, it reduces the matching efficiency. In order to solve the fore-mentioned problems, this study proposes a novel map matching algorithm that combines twice gridding and weighting factors method, acquiring candidate road segments based on grid2, locating the shortest path by grid1. The method inherits the algorithm from the map matching algorithm based on twice gridding; Weighting factors map matching algorithm proposed in this study is improved by dynamically adjusting the factors weights by speed, interval length, etc., and integrates advantages of twice gridding method. The matching result has proven that the matching accuracy rate in specific situations such as parallel road segment, intersection area, intensive road segment or even large local positioning error is improved, and the failure of the driving direction and the trajectory angle at a low speed is avoided. The proposed algorithm outperforms the normal weighting factor algorithm, and it also has better performance in the case of similar road structure, large deviation of positioning and intensive road segment, And improve the matching efficiency by twice gridding method.

The map matching algorithm proposed in this study can provide basic technical support for urban traffic analysis, and also explores a new perspective for the map matching study. Since the data type of the algorithm in this study is offline floating car data, the next step is to handle online floating car data based on big data processing platform such as Hadoop, improve the processing timeliness.

References

1. Hashemi, M.: A critical review of real-time map-matching algorithms: current issues and future directions. *Comput. Environ. Urban Syst.* **48**, 153–165 (2014)
2. Kim, J.: Node based map matching algorithm for car navigation system. In: *International Symposium on Automotive Technology & Automation, Florence. Global deployment of advanced transportation telematics/ITS* (1996)
3. Bernstein, D.: An introduction to map matching for personal navigation assistants. *Geom. Distrib.* **122**(7), 1082–1083 (1998)
4. White, C.E., Bernstein, D., Kornhauser, A.L.: Some map matching algorithms for personal navigation assistants. *Transp. Res. Part C* **8**(1), 91–108 (2000)
5. Taylor, G.: Road reduction filtering for GPS-GIS navigation. *Trans. GIS* **5**(3), 193–207 (2001)
6. Morikawa, T., Miwa, T.: Preliminary analysis on dynamic route choice behavior: using probe-vehicle data. *J. Adv. Transp.* **40**(2), 140–163 (2006)
7. Greenfeld, J.S.: Matching GPS observations to locations on a digital map. In: *Transportation Research Board 81st Annual Meeting* (2002)
8. Quddus, M.A.: A general map matching algorithm for transport telematics applications. *GPS Solutions* **7**(3), 157–167 (2003)
9. Syed, S., Cannon, M.E.: Fuzzy logic based-map matching algorithm for vehicle navigation system in urban canyons. In: *Ion National Technical Meeting* (2004)
10. Ren, M.: A fuzzy logic map matching for wheelchair navigation. *GPS Solutions* **16**(3), 273–282 (2012)
11. Ren, M., Karimi, H.A.: A hidden Markov model-based map-matching algorithm for wheelchair navigation. *J. Navig.* **62**(3), 383–395 (2009)
12. Wu, D.: *Research on the Computational Method based on Markov chain for Road Weight of Traffic Network*. Jilin University (2011)
13. Cao, P.: *Research on map matching algorithm for large scale probe vehicle data*. Tsinghua University (2011)
14. Zhao, S., Zhang, J., Qu, R.: An improved map matching algorithm for floating car. *Bull. Surv. Mapp.* **01**, 97–102 (2018)
15. Ochieng, W.Y., Quddus, M., Noland, R.B.: Map-matching in complex urban road networks. *Revista Brasileira De Cartografia* **55**(2), 1–14 (2003)
16. Fan, N., Yu, Q.-q., Kang, J.: Map matching algorithm for urban road network based on dynamic weight. *Meas. Control Technol.* (2018)



A Study on the Decision-Making Heterogeneity of Parking Mode Choice

X. H. Li, L. X. Wang^(✉), X. H. Sun, and Z. Zuo

School of Civil Engineering and Architecture, Xinjiang University,
Urumqi, Xinjiang, China

lixuhui_xinda@163.com, xjwanglx@foxmail.com,
xhsun347@foxmail.com, zuozhiukk@foxmail.com

Abstract. A variety of models are applied to study the decision-making process of decision-makers under the framework of behavioral decision theory. The fitting precision and power of interpreting reality may differ among models. In the present work, a mixed logit model, a prospect theory model and a random regret minimization model are applied to study parking mode choice behavior. The heterogeneity of decision rules among decision-makers is explored through a case study considering context-dependence. Results show that decision rules are different for different decision-makers in same decision situation, and decision-making rules for same decision-makers in different scenarios are also different. Thus, the decision-making groups can be demarcated according to different decision rules.

Keywords: Parking mode choice · Decision-making heterogeneity · Context-dependence · Group division · Behavioral decision theory

1 Introduction

As a main theoretical method in the research field of mode choice behavior, behavioral decision theory has been developed for decades, from earlier random utility theory to later prospect theory (PT) and regret theory (RT). In the development of behavioral decision theory, PT and RT have addressed some limitations of random utility theory. However, in practical applications, each model has its own advantage in terms of explanatory power.

This study focus on decision-making rules in parking mode choice behavior, which is a major topic in the field of transportation research. Qin [1] established a nested logit model for long-term parking and short-term parking respectively to study parking choice preference of travellers in parking facilities outside the airport. Ibeas [2] studied drivers parking choice among free on-street parking, toll on-street parking and underground parking by establishing a mixed logit model (ML). And the study indicated that vehicle age is an important variable in addition to parking cost. Furthermore, scholars also used a variety of models to explore the driver's parking decision under uncertainty by considering the impacts of decision situations and heterogeneity of travelers preference. Chaniotakis [3] studied driver's parking choice behavior under uncertainty by establishing a multinomial logit model (MNL), nested logit model

(NL) and ML model. The predictive ability of the parking mode choice models would be enhanced when the drivers' preference heterogeneity and the drivers' individual differences are included. Chen [4] studied site choice of parking transfer users under uncertainty. The utility function of a ML model was established by using the procedure of mean variance under the framework of cumulative prospect theory (CPT). Xie [5] established reference points with a single variable, and obtained the prospect value of each reference point by value function and weight function. Then, the multi-objective decision weight assignment method with fuzzy preference was used in determining weight of influencing factors. Finally, a choice model based on prospect theory is set up. Chorus [6] established parking lot choice models based on random regret minimization and random utility maximum, respectively. And compared the results of the parameter calibration.

In the existing literatures, various alternatives and variables are considered in the studies of parking mode choice. Neither might it be shown same preference for different decision-makers on same alternative, nor importance of an attribute, which leads to differences in the analysis of decision rules. There will be a heterogeneity of decision rules among decision-makers. Wu [7] compared the random utility theory with the RT, indicating that decision-makers do have heterogeneity of decision rules when making route choice decisions. Charoniti [8] studied the combined effects of the arrival time variation and choice situation on route choice by establishing a hybrid route choice model under uncertainty. The results showed that when choice conditions were uncertain, choice behavior was context-dependent.

Researchers have explored many aspects of travel choice behavior, and the research results of predecessors have laid a good theoretical foundation for this study. However, when the issue of parking choice behavior was studied, most studies failed to consider the heterogeneity of decision-making rules among decision-making groups. Comparative studies on several models based on situation dependence and group division are relatively rare, and so are studies exploring the mainstream decision rules of different groups in a specific decision-making situation. In order to make up this research gap, the mainstream decision rules of different groups is studied in this paper under the decision-making situation of parking mode choice, which based on random utility theory (RUT), PT and RT respectively. Section 2 includes a brief introduction of basic assumptions and concepts of different behavioral decision-making theories. Section 3 introduces the theoretical model of behavioral decision-making adopted. In Sect. 4, the models are verified by a case study. Finally, our main conclusions are presented in Sect. 5.

2 Introduction to Several Behavioral Decision Theory

The premise of this study is that decision rules are heterogeneous depending on decision-making situations. In reality, there are various decision rules, while this study focus on three widely-used decision rules based respectively on RUT, PT and RT. The following is a brief introduction to three theories.

In 1950s, Von Neumann and Morgenstern proposed behavioral decision theory under uncertainty based on objective probability theory and the assumption of rational

human, namely expected utility theory. Further, McFadden developed expected utility theory into RUT, which is the predominant theory in behavioral decision making. The RUT assumes that decision maker are completely rational, and follow maximum utility rule when make decisions.

Some scholars have found that decision-makers are not completely rational. They may not be able to obtain complete information, and they also do not completely follow maximum utility rule, which counters to the assumptions of RUT. And then some began to consider decision theories based on bounded rational hypothesis, PT proposed by Kahneman and Tversky [9, 10] as well as RT proposed by Bell [11] are the most used-widely theories.

PT holds that decision-maker sets a base point psychologically when making decision, called reference point, which is used to partition the gains and losses, and value and weight function are used to calculate prospect value for each alternative, the alternative with the maximum prospect value was chose; RT holds that decision-maker compares the possible results that produced by their chosen alternatives with those produced by other alternatives, and they will fell regret psychologically if other alternatives will result in better results, otherwise, they will fell rejoice. The alternative with the minimum regret value was chose.

3 Constructing Behavioral Decision Models

3.1 Random Utility Model

MNL model is derived by Luce under the assumption of decision-makers who always follow maximum utility rule, and it is widely used. However, there are three prime limitations: the first one is that it cannot express the change of random preference; the other one is the independence between the irrelevant alternatives; the third one is that it cannot deal with the correlations of unobserved factors in different periods. The ML model not only can deal with three limitations of MNL model, but also approximates any random utility model when random coefficients obey different distributions.

The utility of ML model is divided into two parts: observable item, V_{ni} , and unobservable item, ε_{ni} , then utility can be expressed as:

$$U_{ni} = V_{ni} + \varepsilon_{ni} \quad (1)$$

Where: $V_{ni} = \beta^T X_{ni}$, X_{ni} is a attribute vector, and β^T is a random coefficient vector, which can obey different distributions with parameter θ , such as normal, lognormal, triangular, uniform and S_B . The probability that decision maker n chooses the alternative i can be obtained from Hensher [12]:

$$P_{ni} = \int \left[\frac{\exp(\beta^T X_{ni})}{\sum \exp(\beta^T X_{ni})} \right] f(\beta|\theta) d\beta \quad (2)$$

Where: $f(\beta|\theta)$ is the probability density function of the distribution β obey.

3.2 Prospect Theory Model

In the decision-making process, reference points of different decision-makers are not necessarily same, therefore, personal attributes need to be considered in reference point. According to reference point setting method in travel route choice, $T_{Desired} = (1 + \beta) \sum_{i=1}^n (T_{i,free}/n)$, travel time reference point of route i , $T_{i,desired}$ can be expressed as a weighted average of the free flow travel time $T_{i, free}$, n is total number of alternative routes. β is route parameter. In this study, free flow time $T_{i, free}$ can be replaced by comprehensive attribute value, A_{ni} , of the parking lot. Route parameters β can be replaced with personal attribute parameter $\bar{\beta}_n$, $\beta_n = \beta_n / \max(\beta_1, \beta_2, \dots, \beta_n)$. In order to simplify the calculation, β_n is expressed in the form of linear addition. $\beta_n = \beta_{n1}x_{n1} + \dots + \beta_{ni}x_{ni} = \sum_{i=1}^I \beta_{ni}x_{ni}$, x_{ni} is the i -th personal attribute value of decision maker n , and β_{ni} is parameter, which can be obtained by regression. The heterogeneous reference point can be expressed as:

$$A_{ni,desired} = (1 + \bar{\beta}_n) \sum_{n=1}^N (A_{ni}/N) = (1 + \bar{\beta}_n) \sum_{n=1}^N (\alpha_{cn}C_i + \alpha_{dn}D_i + \alpha_{m}T_i)/N \tag{3}$$

Where: C_i , D_i and T_i are parking fee, walking distance and waiting time, respectively; α_{cn} , α_{dn} and α_m are weight coefficients of parking fee, walking distance, and waiting time, respectively. The value function and weight function used are:

$$v(x) = \begin{cases} x^\alpha & x \geq 0 \\ -\lambda(-x)^\beta & x \leq 0 \end{cases} \quad \pi(p) = \begin{cases} p^\gamma / [p^\gamma + (1-p)^\gamma]^{1/\gamma} & x \geq 0 \\ p^\delta / [p^\delta + (1-p)^\delta]^{1/\delta} & x \leq 0 \end{cases} \tag{4}$$

Where: α and β are risk attitude factors, λ is loss avoidance factor.

The prospect value is expressed as: $V(f) = V(f^+) + V(f^-)$

$$V(f^+) = \sum_{i=0}^n v(x_i)\pi^+(p_i), V(f^-) = \sum_{i=-m}^0 v(x_i)\pi^-(p_i) \tag{5}$$

3.3 Regret Theory Model

The development of RT model has undergone the original Random Regret Minimization [13] (O-RRM), the Classic Random Regret Minimization [7] (C-RRM) and the generalized random regret minimization [14] (G-RRM). Among them, the C-RRM model has a smooth likelihood function compared to the O-RRM model, and has fewer parameters than the G-RRM model. The specific C-RRM model is expressed as:

$$RR_i = \sum_{j \neq i} \sum_m \ln\{1 + \exp[\beta_m(x_{jm} - x_{im})]\} + \varepsilon_i \tag{6}$$

4 Experiment

In order to study behavioral decision theory of situational dependence and decision-making group dependence under specific application conditions, the authors conducted a questionnaire survey on parking mode choice in Urumqi. In the questionnaire, two types of parking, surface parking and garage parking, were set up for respondents. The survey consists of three parts, the first part is the personal attributes of decision maker; the second part is choice scenarios; the third part is data survey on calibrating prospect theory value function parameters. 466 valid questionnaires were collected by online and field survey, and 3502 ML model samples, 2382 PT model samples, and 1751 C-RRM model samples were obtained. Based on the survey data, ML model, PT model and RT model were established and estimated. In the ML model, personal attribute variables are set in the form of dummy variables, variables used in PT model are same as ML's, C-RRM model used in this study does not involve personal attributes, therefore, parking lot attributes are selected as variable of this model, and variable definitions are shown in Table 1.

Table 1. Variable definitions.

Influence factor category	Influence factors		Variables
Personal attribute	Gender	1: male; 0: female	Gender
	Occupation	1: institution; 0: other	Occ1
		1: enterprise; 0: other	Occ2
		1: individual; 0: other	Occ3
		1: freelancers; 0: other	Occ4
	Annual income	1: 50–80 K; 0: < 50 K	Inc1
		1: 80–120 K; 0: < 50 K	Inc2
1: 120–200 K; 0: < 50 K		Inc3	
1: > 200 K; 0: < 50 K		Inc4	
Parking property	Parking rate	6: 6yuan/h; 5: 5yuan/h; 4: 4yuan/h; 3: 3yuan/h; 2: 2yuan/h	Cost
	Walking distance	2: < 200 m; 3.5: 200–500 m; 5: > 500 m	Distance
	Waiting time	5: < 5 min; 10: 5–15 min; 15: > 15 min	Waitim

The estimation results of ML model are shown in Table 2. Table 3 shows the estimated parameters of the C-RRM model, B_PC, B_PD, and B_WT are perception parameters of parking rate, walking distance and waiting time, respectively. For PT model, the value function parameter is recalibrated according to the least squares proposed by Xu [15], and final calibration results are $\alpha = 0.43$, $\beta = 0.36$, $\lambda = 1.43$.

About estimation results, positive coefficient indicates that variable has a positive effect on alternatives, and vice versa, coefficient value indicates the influencing degree. It can be seen that P-values of Gender, Occ1, Occ2, Inc1, and Inc2 are less than 0.05, which indicate these variables affect significantly parking mode choice.

Table 2. ML model estimation results.

Var	Coef.	Std. Err.	Z	P > z	[95% Conf. Interval]	
Gender	-0.293	0.141	-2.08	0.038	-0.570	-0.016
Occ1	1.114	0.376	2.96	0.003	0.377	1.852
Occ2	0.889	0.353	2.53	0.011	0.200	1.579
Occ3	0.290	0.642	0.45	0.651	-0.970	1.548
Occ4	0.796	0.587	1.36	0.174	-0.353	1.947
Inc1	1.172	0.479	2.45	0.014	0.234	2.111
Inc2	0.937	0.447	2.10	0.036	0.061	1.814
Inc3	0.524	0.486	1.08	0.280	-0.427	1.477
Inc4	0.428	0.572	0.75	0.454	-0.693	1.550
Cost	-0.430	0.210	-2.05	0.040	-0.841	-0.019
Waitim	-0.231	0.058	-3.95	0.000	-0.346	-0.117
Distance	-0.374	0.159	-2.36	0.018	-0.685	-0.064
Log likelihood: -906.02457						
LR chi2(3): 205.15						
Prob > chi2: 0.000						

Table 3. RT model estimation results.

Var	Coef.	Std err	t-test	p-value	Robust Std err	Robust t-test	p-value
B_PC	-0.379	0.051	-7.46	0.00	0.050	-7.63	0.00
B_PD	-0.096	0.052	-1.85	0.06	0.051	-1.87	0.00
B_WT	-0.143	0.011	-13.45	0.00	0.011	-13.21	0.00
Likelihood ratio test: 381.306							
Rho-square: 0.157							
Adjusted rho-square: 0.155							

According to the meaning of context-dependence, different decision-makers have different decision rules in different scenarios, different decision-makers refer to groups of decision makers with different personal attributes, and different scenarios refer to different selection situations. This paper explores the situation dependence from two perspectives: one is decision-making rules of same decision-makers category in different decision scenarios (Fig. 1 - left), and the other is decision rules of different decision-makers categories in same scenario (Fig. 1 - right).

From models' estimation results, variables with insignificant effects are rejected, and decision-makers are divided into eight categories according to different levels of variables with significant effects. Based on classifications, predicted and actual proportions of different decision-makers categories in same scenario are shown in Table 4. For the same category, it is obvious that predicted proportions of each model are significantly different from the actual proportion. From the perspective of male, predicted proportions are relatively close to actual choice results, while that of female

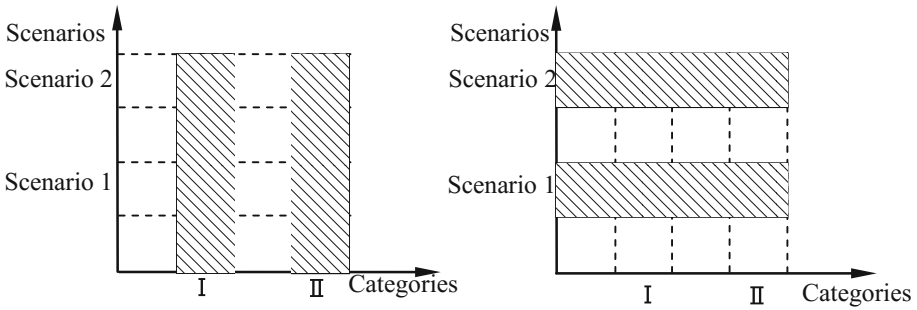


Fig. 1. Scenario-dependent research classifications

Table 4. Predicted and actual proportions of garage parking for different decision-makers categories.

Personal attribute			ML model		C-RRM model		PT model	
			Actual	Predict	Actual	Predict	Actual	Predict
Male	Institution	50–80 K	29.6%	33.8%	29.6%	26.6%	70.2%	46.8%
		80–120 K	46.4%	42.0%	46.4%	27.1%	63.3%	59.1%
	Enterprise	50–80 K	59.0%	51.4%	59.0%	26.3%	65.4%	67.9%
		80–120 K	50.7%	55.0%	50.7%	27.1%	67.2%	58.8%
Female	Institution	50–80 K	45.7%	50.0%	45.7%	27.8%	70.0%	42.5%
		80–120 K	47.3%	54.5%	47.3%	27.4%	73.7%	52.6%
	Enterprise	50–80 K	69.8%	62.8%	69.8%	28.3%	64.9%	40.5%
		80–120 K	38.7%	35.5%	38.7%	25.3%	65.1%	62.7%

decision-makers appear to be more inconsistent. The reason for this result may be that most females are relatively emotional, most of the time. On the basis of classification divided by incomes and occupations, decision-makers’ decision-making results are diversified, and don’t bias toward certain decision rule.

Predicted and actual choice proportions of each model for same decision-maker in different scenarios are shown in Table 5. Among four different scenarios, the predicted results are the closest to actual proportions in scenario 3, while predicted proportions differ greatly from actual proportions in other scenarios.

Table 5. Proportions of garage parking choice in different scenarios.

Scene number	Actual	ML Predicted	C-RRM Predicted	PT Actual	PT Predicted
1	82.6%	49.0%	36.0%	73.5%	54.3%
2	75.8%	48.3%	43.8%	72.0%	52.1%
3	48.7%	50.9%	49.8%	39.0%	37.5%
4	65.8%	48.3%	43.8%	73.5%	49.5%

It should be noted that the questionnaire on PT is different from that on RUM and RT, in order to ensure comparability of predictions, Table 5 only shows the results of scenarios with same parking lot attribute levels.

In this study, the deviation between predicted proportions and actual proportions for each model is measured by the method of absolute error value, which can be expressed as $R_{ML,i} = |A_{ML,i} - P_{ML,i}|$, i is different decision-makers categories (for Table 6) or different scenarios (for Table 7), $A_{ML,i}$ and $P_{ML,i}$ are actual and predicted proportion of ML model, respectively. When determine decision-making rule, we believe that decision-makers will follow the decision rule corresponding to a model with minimum deviation, and it is treated as the mainstream decision rule, which can be expressed as $DR_i = \min\{R_{ML,i}, R_{C-RRM,i}, R_{PT,i}\}$, DR_i is decision rule. The deviations of each model and mainstream decision rules for each decision-maker category or each scenario are shown in Tables 6 and 7, respectively.

Table 6. Mainstream decision rule for each decision-makers category.

Personal attribute			ML error	C-RRM error	PT error	Decision rule
Gen	Occ	Inc				
Male	Institution	50–80 K	0.042	0.030	0.234	Minimum regret
		80–120 K	0.044	0.193	0.042	Maximum prospect
	Enterprise	50–80 K	0.076	0.327	0.025	Maximum prospect
		80–120 K	0.043	0.236	0.084	Maximum utility
Female	Institution	50–80 K	0.043	0.179	0.275	Maximum utility
		80–120 K	0.072	0.199	0.211	Maximum utility
	Enterprise	50–80 K	0.070	0.415	0.244	Maximum utility
		80–120 K	0.032	0.134	0.024	Maximum prospect

Table 7. Mainstream decision rule for each scenario.

Scene number	ML error	C-RRM error	PT error	Decision rule
1	0.336	0.466	0.192	Maximum prospect
2	0.275	0.320	0.199	Maximum prospect
3	0.022	0.011	0.015	Minimum regret
4	0.175	0.220	0.240	Maximum utility

From statistical results in Table 6, it can be seen that there are significant differences in mainstream decision rules for different decision-makers groups when making decisions in same decision scenario. The decision rule followed by female decision-makers who are institutions staffs is maximum utility, that followed by male decision-makers who are institutions staffs with annual incomes of 50–80 K is minimum regret, however, male decision-maker who are institution staffs with annual income of 80–120 K follow the decision rule of maximum prospect. Among the enterprise personnel, decision rules followed by women who have annual income of 50–80 K is

maximum utility, that followed by women whose income is 80–120 K is maximum prospect. Decision-making groups can be divided according to different decision rules that different decision-makers followed in the same scenario.

From statistical results in Table 7, it can be seen that decision rules followed by same decision-makers category in different scenarios are also different, indicating that decision-makers are context-dependence when making decisions, which may be related to the personal attributes of decision-makers. Comparing with institution staffs, enterprise personels pay more attention to the time; Comparing with people whose income are 50–80 K, people who have an annual income of 80–120 K have less concern on cost; male and female decision-makers may have different requirements for walking distance.

5 Conclusion

Parking mode choice model research is a common concern of traffic managers and traffic researchers. Previous studies usually assume that decision-makers only use a single decision rule, and rarely consider decision-rule heterogeneity between decision-makers groups. In fact, different decision-makers groups tend to follow different decision rules depending on different decision-making situations.

We take the heterogeneity of decision rules between decision-maker groups as premise, and apply three widely-used decision-making theories on survey data to analyze and model parking mode choice behavior, aiming to explore heterogeneity of decision rules followed by decision-makers. The models estimation results show following conclusions:

- (1) Different decision-makers follow different decision rules when making decisions in same scenario, and the same decision-maker group has different decision rules in different scenarios, that is, decision-makers are context-dependent when making decisions.
- (2) According to the characteristic of decision-makers are context-dependence when making decisions, decision-makers can be divided into different groups and to which a certain choice model can be applied.
- (3) It can be known from Table 6 that when the decision-maker's gender and occupation are same (male, institutional) and the income is different, decision rules followed by decision-maker are different; therefore, accurately dividing income level is beneficial to model accuracy.

This study verifies the existence of decision-rule heterogeneity between different decision-makers groups, but there are still some problems need to be further studied. At present, it is assumed that each decision-maker group follow one single decision rule, but the reality may not be same as this case. Whether the decision rule for the same decision-maker group is single or mixed decision rules needs to be further explored. If it is a mixed decision rule, how the various decision rules affect on the final choice results still needs to be studied in depth.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

References

1. Qin, H.M.: Nested logit model formation to analyze airport parking behavior based on stated preference survey studies. *J. Air Transp. Manage.* **58**, 164–175 (2017)
2. Ibeas, A.: Modelling parking choices considering user heterogeneity. *Transp. Res. Part A* **70**, 41–49 (2014)
3. Chaniotakis, E.: Drivers' parking location choice under uncertain parking availability and search times: a stated preference experiment. *Transp. Res. Part A* **82**, 228–239 (2015)
4. Chen, C.M.: Influence of parking on train station choice under uncertainty for park-and-ride users. *Procedia Manufact.* **3**, 5126–5133 (2015)
5. Xie, B.L.: Parking mode choice based on prospect theory. *J. Trans. Sci. Eng.* **28**(3), 71–76 (2012)
6. Chorus, C.G.: A new model of random regret minimization. *Eur. J. Transp. Infrastruct. Res.* **10**(2), 181–196 (2010)
7. Wu, J.M.: A hybrid discrete choice model with heterogeneous decision rules on attributes and research on stochastic. Nanjing University (2017)
8. Charoniti, E., Rasouli, S.: Context dependent latent class behavioral mixture model of utility maximization and regret minimization decision-making under uncertainty. TRB 2017 Annual Meeting
9. Kahneman, D.: Prospect theory: an analysis of decision under risk. *Econometric Rev.* **47**(2), 263–292 (1979)
10. Tversky, A.: Advances in prospect theory cumulative representation of uncertainty. *J. Risk Uncertainty* **5**, 297–323 (1992)
11. Bell, D.E.: Regret in decision making under uncertainty. *Oper. Res.* **30**(5), 961–981 (1982)
12. Hensher, D.A.: The mixed logit model: the state of practice. *Transportation* **30**, 133–176 (2003)
13. Chorus, C.G.: A random regret-minimization model of travel choice. *Transp. Res. Part B* **42**, 1–18 (2008)
14. Chorus, C.G.: A generalized random regret minimization model. *Transp. Res. Part B* **68**, 224–238 (2014)
15. Xu, H.L.: A decision-making rule for modeling travelers' route choice behavior based on cumulative prospect theory. *Transp. Res. Part C* **19**, 218–228 (2011)



An Improved Vehicle Detection and Tracking Model

Libin Hu^{1,2}, Zhongtao Li^{1,2}(✉), Hao Xu^{1,2}, and Bei Fang^{1,2}

¹ School of Information Science and Engineering, University of Jinan,
Jinan 250022, China

hlblydx@163.com, ise_lizt@ujn.edu.cn,
944507073@qq.com, ll3847750@qq.com

² Shandong Provincial Key Laboratory of Network Based Intelligent
Computing, University of Jinan, Jinan 250022, China

Abstract. With the continuous advancement of urbanization in China, the number of urban motor vehicles has increased exponentially, but the pressure on urban traffic and security monitoring has also increased. In recent years, the application of Intelligent Transportation System (ITS) has created good economic benefits for the transportation industry. Vehicle Detection and Tracking is the basis for subsequent vehicle information attribute calculation and statistical analysis of data and plays a leading role in ITS. This paper discusses the state-of-the-art detection and tracking algorithms and points out their shortcomings, and on this basis, we proposed an improved video vehicle detection and tracking model: a model based on Gaussian mixture model, Kalman filter + CAMshift + contour method. First, in the background extraction process, the Gaussian mixture model with different parameters is used for the foreground and the background to reduce the calculation cost and enhance the separation effect; For the extracted foreground, the part connecting in different vehicles is overlapped by the matching of the pits to distinguish the overlapping vehicles; In the tracking process, the results of CAMshift and contour tracking are used as observations of the Kalman filter to avoid the phenomenon that vehicles with similar color distribution characteristics and similar positions are misidentified as the same target. The experimental results show that our model has better robustness to complex road conditions, and the detection and tracking effect is better than other models.

Keywords: Video vehicle detection and tracking · Mixed gaussian model · Overlapping segmentation · Weighted CAMshift

1 Introduction and Related Algorithm

Video vehicle detection and tracking is a prerequisite for vehicle speed calculation, road congestion detection and license plate recognition, and is an important part of ITS. At present, there are many methods for detecting and tracking vehicles based on video,

The original version of this chapter was revised: The corrections in author affiliation have been incorporated. The correction to this chapter is available at https://doi.org/10.1007/978-981-13-7542-2_29

© Springer Nature Singapore Pte Ltd. 2019
X. Zeng et al. (Eds.): ITASC 2019, SIST 127, pp. 84–93, 2019.
https://doi.org/10.1007/978-981-13-7542-2_8

the processes of detection and tracking are basically same. The video is preprocessed (generally de-noising) to avoid various interferences, and then moving vehicles in the image sequence are extracted by different methods. The morphological processing was performed (generally dilation processing) on the extracted foreground target, and then uses the tracking algorithm to track the detected moving vehicle. Among them, the noise removal processing stage mainly uses mean filtering, median filtering, Gaussian filtering, and bilateral filtering. Due to different sources of noise, the noise mainly includes salt and pepper noise, Gaussian noise, impulse noise, and Poisson noise. We select the corresponding filter to remove noise according to different noise types.

1.1 Filters and Usage Scenarios

The filter is classified into a linear filter and a nonlinear filter according to a mathematical operation or a logical relationship in principle. The previously mentioned mean filtering, Gaussian filtering are linear filter, and median filtering, bilateral filtering are nonlinear filter.

Mean filtering also known as neighborhood averaging filtering, it replaces each pixel value in the image with the mean of all other pixel values in its local neighborhood, and we can get a smoother image. The mathematical expression of the mean filtering is:

$$I(x, y) = \frac{1}{M} \sum_{(k,l) \in N} I(k, l)$$

Where $I(x, y)$ represents the pixel value at the center of the image at point (x, y) . M represents the total number of pixels in the neighborhood N of the pixel. The size of the neighborhood N determines the degree of this mean filtering, and the filter with large neighborhood can remove large noise.

Gaussian filtering selects different weights according to the different distribution of pixels in the image. It is suitable for noise types that conform to the normal distribution. The Gaussian function in a two-dimensional image is:

$$g(x, y) = e^{-\frac{(x^2 + y^2)}{2\sigma^2}}$$

The parameter variance- σ determines the width of the Gaussian distribution. The median filtering uses the median value of the neighboring pixel points of current pixel to replace the gray value of current. It does not depend on the point that is too different from the neighborhood pixel, so it can remove the impulse noise and the salt and pepper noise well, it also is the method used in our work when dealing with noise. Bilateral filtering is a relatively in-depth de-noising method studied and developed in recent years. We know that bilateral filtering is sensitive to different kinds of weak noise and can preserve the original geometric features of the image. Bilateral filtering is a kind of compromise processing that combines the spatial proximity of images and the similarity of pixel values. It has the characteristics of considering spatial information and gray similarity at the same time, and achieves the purpose of retaining edges and removing noise.

$$g(i, j) = \frac{\sum_{k,l} f(k, l) w(i, j, k, l)}{\sum_{k,l} w(i, j, k, l)}$$

The output of the bilateral filter is composed of two parts, that is, the weight parameter $w(i, j, k, l)$ is obtained by the product of the geometric space distance coefficient and the pixel difference coefficient.

1.2 Vehicle Detection Algorithms

With the development of machine vision, analyzing and understanding an image sequences has become a hot research field of computer vision. Extracting moving targets from continuous images is a prerequisite for our follow-up work. The vehicle detections discussed in this paper are based on a sequence of sequential images acquired by a static camera. Specifically, in video vehicle detection, common detection algorithms include frame difference method, optical flow method, background difference method, and detection line method. In our work, the improved Gaussian mixture model based on background subtraction is used to extract moving vehicles.

The frame difference method is to detect a moving targets by using a moving object to change a pixel in a continuous video sequence, and the moving targets can be extracted by comparing the difference between successive images of the target and the artificially set threshold. Supposing $f(x, y, t)$, $f(x, y, t + 1)$ respectively represent the pixel values of the image at the point (x, y) at the t and $t + 1$ moments in the image sequence, and the difference image $D(x, y)$ can be expressed as:

$$D(x, y) = |f(x, y, t_i) - f(x, y, t_j)|$$

Then All pixels are binarization by a threshold T set in advance. In the preprocessing of the frame difference method, after de-noising, the color map is generally converted into a gray image to reduce the amount of calculation, and then the current image is compared with the image of the previous frame, and then the binarized difference image is our output image, some morphological operations are appropriately taken according to the situation of output image, and the most is the dilation operation, filling small holes and non-connected regions to achieve a better foreground segmentation effect. The frame difference method has a small amount of calculation, good real-time performance, and good robustness to illumination changes, but it is sensitive to the speed of the vehicle, and the difference between the vehicles of different speeds is large.

The optical flow method changes the image of the detection area into a vector field of velocity, and each vector represents an instantaneous change in the position of a point in the image. If there is no moving object in the image, the optical flow vector corresponding to each pixel point changes continuously and uniformly. When there is relative motion between the vehicle and the background, the vector change of the vehicle must be different from the vector change of the background, so the moving vehicle is detected based on the uneven change in the velocity vector. In video vehicle detection, the gradation of the optical flow field is not conserved due to multiple light sources, noise, etc., because a relatively accurate velocity field cannot be got, and

iterations require multiple cycles to converge, which cannot satisfy the Real-time requirements for video vehicle detection.

The background difference method is a widely used vehicle detection method in the current intelligent transportation system. The background is extracted from the changed image sequence, and then the sequence image is compared with the background image to obtain a foreground image containing the moving object. Similarly, the background difference method also compares the result of subtracting the original image from the background image with the set threshold. If the threshold is exceeded, it is considered to be a moving vehicle in the image scene, wherein the background image extraction process generally requires the first few frames, the few images is trained to fit the optimal result. The background difference method is as follows:

$$D_K(x, y) = |f_k(x, y) - B_k(x, y)|$$

$f_k(x, y)$ refers to the brightness value of the pixel at the (x, y) position at time k , $B_k(x, y)$ is the brightness value of the background model, and the result is the difference image after subtraction. The same result $D_k(x, y)$ is compared to the threshold T , and the pixel position greater than T is considered to be a moving vehicle. The line detection method is also the virtual coil method. The principle is similar to the underground sensing coil, which is a straight line perpendicular to the direction of travel of the vehicle. We use the background modeling algorithm to obtain the background image or the edge detection image on the detection line when there is a moving vehicle. When the line is detected and the width of the vehicle coverage detection line is greater than the set threshold, it is considered that a moving vehicle is detected.

There is also a model-based vehicle detection algorithm. For example, the 3D model method is used to identify and track moving vehicles. Then we use the vehicle identification method to identify and track. Generally, we think that the 3D model method combines various sports vehicle features, such as color distribution, velocity vector edge, etc., which is more comprehensive and scientific. But the hardware computing resources of modeling method are high, which is not conducive to real-time detection, and the probability that the vehicle with the same artificial setting feature could have a incorrect detection. The mixed Gaussian model method [5] is the main vehicle detection method discussed and studied in our work. Obviously, the single Gaussian model cannot fully describe the probability distribution of the background. In the background training process, the mixed Gaussian model first initializes the background with the maximum expectation-EM algorithm, and then classifies the feature values of each pixel into K . One of the Gaussian distributions, which should include foreground and background distributions, using different model results to divide the image into foreground and background. Background extraction and background updating are included in the background extraction process, and background initialization has an important influence on the speed and quality of background modeling. When using Gaussian algorithm, if the learning rate is small, it takes a long time to remove the foreground vehicle. In addition, mean filtering and median filtering are often used to initialize the background. The background update method takes the mean filter update background as an example. First, the background model pixel value at the point (x, y) is obtained from an image $F(x, y, i)$:

$$B(x, y, t) = \frac{\sum_{i=1}^n F(x, y, i)}{N}$$

The difference image of the current frame and the N th frame background model is calculated by using the $N + 1$ frame image, and the difference image is binarized according to the set threshold value, and the result of this time update is obtained. According to experience, we should judge whether the current pixel in the previous frame belongs to the background or foreground before updating the background. We update the pixel when it is the background. If it is not the background, it will not be updated. This can avoid the situation of high-speed vehicles with ghosting.

1.3 Vehicle Tracking Algorithms

Tracking the vehicles detected in the video is the basis for getting the speed, traffic flow, and so on. Vehicle tracking is to detect the moving vehicle and process the target and obtain its motion track. After detection and tracking, feature extraction can be performed on the moving vehicle to obtain various attributes of the vehicle, such as quantity, speed, license plate recognition, and the like. At present, video-based traveling vehicle tracking methods can be divided into model-based tracking, contour-based tracking, feature-based tracking, and tracking based on motion state prediction. In our work, the feature-based CAMShift + contour tracking weighting method is used to obtain the vehicle position result of the next frame, and this result is taken as the observation value of the Kalman filter, combined with the predicted value of the Kalman filter to finally obtain the vehicle of the next frame position.

The predictive tracking algorithm based on vehicle moving state is a hot topic in recent times. Combining the filtering theory to remove the noise in the image and predicting the position of the vehicle target to obtain the predicted value at the next moment, and then predicting the predicted value based on the actual measured result. Most of which are closer to the predicted value and the actual measurement are worth weighting. Kalman filtering and particle filtering are now the most used in target trajectory tracking. Kalman filtering is a linear system state equation. The observation data is input and output through the system. Since the observation data contains noise, the optimal estimation is equivalent to the de-noising process. The Kalman filtering assumes that both the predicted noise and the measured noise in the distribution of Gaussian, the weighted filter result for the predicted value and the measured value is used as the final positioning result. The weight value depends on the degree of uncertainty from measurement process and the prediction process. Compared to the Kalman filter, the particle filters (PF) do not make any assumptions. The Monte Carlo method is directly used to represent the probability by the particle set. Its fundamental thought is to express its distribution through random status particles extracted from posterior probabilities.

The contour-based and motion-based tracking methods use the vehicle as the smallest unit of matching, but feature-based tracking is the use of some or some combination of features of the vehicle as the smallest element of the match. This kind

of algorithm extracts features from each vehicle, such as distinct lines or corners, color histograms, textures, centroids and etc. The features used can be a distinct feature or a combination of several features to increase the matching accuracy. The algorithm is still effective in detecting partially occluded vehicles, because some features of partially occluded vehicles are still visible. Feature-based tracking method just makes up for the shortcomings of contour-based method, so we adopts the weighted method of both. The well-known continuous adaptive mean shift (CAMShift) algorithm belongs to the feature-based tracking method. The method is based on the classical Mean-shift. But the Mean-shift method is only based on the color histogram of the moving target, so CAMShift is same. For the single feature tracking method, it uses the mean shift method in each frame image to initialize the next frame image using the position result of the current frame, thereby completing the tracking of the vehicle in the continuous image.

2 Related Work

As early as the 1960s and 1970s, the United States, Japan and Europe began research on traffic monitoring systems. Since the introduction of the concept of ITS in the 1990s, research on intelligent transportation systems has become increasingly hot. The development of China's intelligent transportation system started late. In the 1990s, we began to pay attention to the development of international ITS and began to invest in research.

In terms of vehicle detection and tracking of intelligent transportation, research results at home and abroad are continuously released and put into the urban road traffic management system. In the aspect of vehicle detection, the paper [1] proposed a global optimal threshold selection algorithm for the two-dimensional least squares method, which makes the background model more balanced and robust. Paper [2] proposes a local threshold method for shape decomposition of background image binarization. Paper [3] proposed a new approach for efficient detection and tracking of moving vehicles in the presence of shadow and partial occlusion under complex road scenes by combining the modified background subtraction and innovative adaptive search window methods. The use of variable-size window (Bounding Box) reduces the searching space proportional to the allocated time of the moving vehicles. Paper [4] adopted a parametric 3D model to describe various types of vehicles in traffic scenes. Paper [5] considered the problem of vehicle video detection and tracking. A solution based on the partitioning a video into blocks of equal length and detecting objects in the first and last frames of the block is proposed. The average tracking accuracy of all vehicles makes about 70%. In order to solve the problems of large computational complexity and large detection delay in the current commonly used traffic congestion detection methods, the paper [6] proposed a road traffic congestion detection algorithm based on image texture analysis to meet the actual needs of traffic management and traffic guidance systems. In order to solve the problem that the traditional Gaussian mixture model has poor adaptive ability to change the moving target speed, paper [7] proposes an improved Gaussian mixture model and Kalman filter vehicle detection and tracking method, which defines the moving target speed factor.

In terms of vehicle tracking, paper [8] uses the vehicle centroid of the Gaussian mixture model stage as the measured value of the Kalman filter to achieve the tracking of the specified target vehicle. In, paper [9] according to the characteristics of strong correlation between frames of dynamic sequence images, combined with the inertia of object kinematics, a dynamic image vehicle recognition and tracking algorithm based on dynamic template and position prediction principle is proposed.

In paper [10], the Kalman filter needs to make assumptions about the noise characteristics and the motion law of the target. A moving target tracking algorithm based on the gray prediction model $GM(1, 1)$ is proposed. The algorithm mines the current motion law of the target through the continuously updated gray prediction model $GM(1, 1)$, which overcomes the shortcomings of the Kalman filter tracking scheme that need to make assumptions when the motion law is unclear and the noise characteristics are not clear. In paper [11], a tracking algorithm based on geometric active contour model and particle filter is proposed, which tracks the deformation target in complex environment and achieves good results.

3 Improved Models for ITS

In the extrusion of previous studies, we proposes an improved video vehicle detection and tracking model. First, in order to improve the real-time performance of the vehicle detection phase, different learning rates σ are improved for the number K of the Gaussian model in the mixed Gaussian model and the foreground background of the velocity; The second, the overlap of the pit matching is performed because of the camera angle and the vehicle overlap caused by other environments, and the phenomenon that the overlapping vehicle is detected as one vehicle is removed. Thirdly, in the vehicle tracking phase, because CAMShift uses a single color feature for tracking, we use the weighted result of CAMShift and contour tracking combined with Kalman filter to achieve vehicle tracking and improve tracking accuracy. The entire detection and tracking model is shown in Fig. 1:

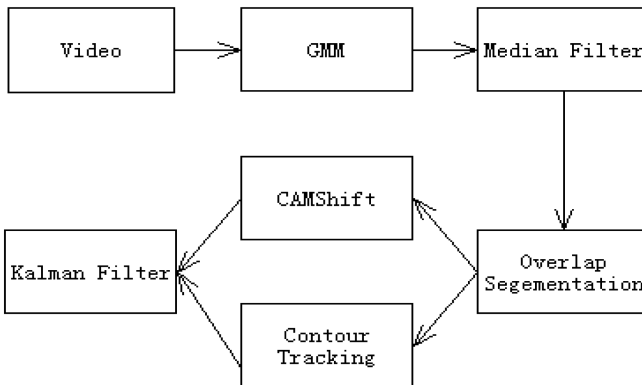


Fig. 1. Vehicle detection and tracking model

3.1 Vehicle Detection

A large number of researches and applications have proved that the hybrid Gaussian model has a good effect on vehicle detection in complex road conditions. The background extracted from video is close to the real situation. The mixed Gaussian model generally uses three to five Gaussian models to adapt to different road conditions. For static camera monitoring, the background is usually static, and it is unreasonable to apply multiple Gaussian models to the stationary background portion. In our model, we first divide the image pixels into static pixels and dynamically changing pixels. We use less Gaussian distribution to describe the static pixels, and more Gaussian distributions for foreground pixels. At the same time, the Gaussian model of the static pixel selects a small learning rate to ensure the stability of the background. The dynamic pixel part first determines whether it belongs to the background, the background uses a smaller learning rate, and the foreground uses a larger learning rate. So that the pixel can occupy a larger weight.

In our work, the first 50 frames of the video are selected to update the background. We used 3 Gaussian filters, the background learning rate is set to 0.001, and the foreground learning rate is set to 0.01. The effect on video foreground and background segmentation is shown in Fig. 2. Due to the illumination and camera shake, there are a lot of salt and pepper noise in the video. After the previous analysis, combined with the type of noise, we used a median filter with a template size of 5 to remove these noises. Due to the angle of the camera, there is an overlap between the foreground vehicles, so that two cars that overlap in the detection are detected as the same target. In [12], the edge contour checking method based on Freeman chain code is proposed, and then the judgment of overlapping vehicles is carried out according to the feature points and duty ratio. In order to improve the real-time detection, this paper uses the overlap segmentation based on pit matching to divide the overlapping part into two targets.



Fig. 2. The background and for foreground

3.2 Vehicle Tracking

In this paper, the CAMShift and Kalman filtering methods are used to achieve tracking, and the contour tracking method is added. Because CAMShift only uses the color features to update the position, when two vehicles with similar color features and similar positions appear, the tracking will have errors. The CAMShift method is very robust to occluded vehicles, and the contour tracking method is not sensitive to color and occlusion for contour detection. This paper proposes a method of weighting the CAMShift method and contour tracking method. After obtaining the centroid position of the maximum value of the probability density function by using the CAMShift method, the centroid of the current target contour method is calculated, and the results of the two are weighted and summed as the predicted value of the Kalman filter, and then the Kalman filter iteratively to obtain the final position by our weighted result and its own measurement result obtained the final position result.

Supposing $F(x, y)$ is the result of CAMShift calculation, and $G(x, y)$ is the result of contour tracking. We calculate the feature similarity λ of the color histogram of all targets in the current image. When the value of λ is greater than 0.7, We set the result weight of CAMShift to 0.6 and the result weight of the contour method to 0.4. When the value of λ does not exist is greater than 0.7, we set the result weight of CAMShift to 0.8 and the result weight of the contour method to 0.2. Our final position result can be express as:

$$\begin{cases} C(x, y) = 0.6 \times F(x, y) + 0.4 \times G(x, y) \lambda_{\max} \geq 0.7 \\ C(x, y) = 0.8 \times F(x, y) + 0.2 \times G(x, y) \lambda_{\max} < 0.7 \end{cases}$$

Finally, using the detection and tracking model of this paper, our detection and tracking effect on the vehicle is shown in Fig. 3:



Fig. 3. Result of our model. We randomly extract three images from the video sequence, which are 2334 frame, and 2447 frame. From the results, we can see that the location where the foreground vehicles are connected has a good segmentation effect, and the position detection contour of each vehicle is reduced, indicating that the tracking result is more accurate.

4 Conclusion and Evaluation

This paper summarizes the existing vehicle detection and tracking methods, and based on those art-of-date methods we proposed an improved model. The improved model achieves good detection and tracking effects while reducing hardware computing resources. However, there are still a lot of problems and deficiencies in the whole model. For example, in the process of separating the background of the Gaussian mixture model, it is better to increase the type of Gaussian model distribution and require less samples to update the background, Is our method better able to achieve detection and tracking effects in the case of a vehicle with shadows.

References

1. Chen, Y.: *Service-Oriented Computing and System Integration: Software, IoT, Big Data, and AI as Services*, 6th edn. Kendall Hunt Publishing (2018)
2. Chen, Y., Hualiang, H.: Internet of intelligent things and robot as a service. *Simul. Modell. Pract. Theory* **34**, 159–171 (2013)
3. Pauer, G.: Development potentials and strategic objectives of intelligent transport systems improving road safety. *Transp. Telecommun. J.* **18**(1) (2017)
4. Oskarski, J., Jamroz, K.: Reliability and safety as an objective of intelligent transport systems in urban areas. *J. KONBiN*, **34**(1) (2015)
5. Abutaleb, A.S.: Automatic thresholding of gray level pictures using two-dimensional entropy. *Comput. Vis. Graph. Image Process.* **47**(2), 22–32 (1989)
6. Park, Y.: Shape resolving local thresholding for object detection. *Pattern Recogn. Lett.* **22** (5), 883–890 (2001)
7. Piao, S., Sutjaritvorakul, T.: Compact data association in multiple object tracking: pedestrian tracking on mobile vehicle as case study. *IFAC PapersOnLine*, **49**(15) (2016)
8. Mayyas, A.R., Kumar, S., Pisu, P., Rios, J., Jethani, P.: Model-based design validation for advanced energy management strategies for electrified hybrid power trains using innovative vehicle hardware in the loop (VHIL) approach. *Appl. Energy*, **204** (2017)
9. Hadi, R.A., George, L.E., Mohammed, M.J.: A computationally economic novel approach for real-time moving multi-vehicle detection and tracking toward efficient traffic surveillance. *Arab. J. Sci. Eng.* **42**(2) (2017)
10. Koller, Y.D., Daniilidis, Y.K., Nagel, Y.Z.H., et al.: Model based object tracking in monocular image sequences of road traffic scenes. *Int. J. Comput. Vis.* **10**(3), 257–281 (1993)
11. Stauffer, C., Grimson, W.: Adaptive background mixture models for real-time tracking. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, pp. 246–252 (1999)
12. Kustikova, V.D., Gergel, V.P.: Vehicle video detection and tracking quality analysis. *Pattern Recognition and Image Analysis*, **26**(1) (2016)
13. Yuan, B., Zhang, Y.: Traffic congestion detection algorithm based on image texture analysis. *J. Shanghai Ship Shipp. Res. Inst.* **38**(4) (2015)
14. Zhou, S.K., Chellappa, R., Moghaddam, B.: Visual tracking and recognition using appearance adaptive models in particle filters. *IEEE Trans. Image Process.* **13**(11), 1491–1506 (2004)
15. Rathi, Y., Vaswani, N., Tannenbaum, A., et al.: Tracking deforming objects using particle filtering for geometric active contours. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(8), 1470–1475 (2007)



Determination of Best Foaming Parameters of Modified Asphalt Using Response Surface Analysis Method

Wen Xiaobo¹(✉), Wu Hao¹, Zhang Peng¹, and Ding Fan²

¹ JSTI Group, Nanjing 211112, Jiangsu, China
wxbl68@jsti.com

² Southeast University, Nanjing 211189, Jiangsu, China

Abstract. In order to explore the foaming principle of SBS modified asphalt, three parameters including asphalt temperature, water consumption and water temperature, were selected to study the foaming effect of modified asphalt. After the experimental data were analyzed, the second-order mathematical model was fitted under the interaction of different foaming parameters. Response surface analysis was adopted when variance analysis met the requirement of significant level, and three-dimensional response surface analysis diagram was drawn to optimize the foaming parameters for the best foaming effect.

Keywords: Foamed asphalt · Response surface analysis · Foaming parameter · Mathematical model

1 Introduction

In recent years, warm mixing technology of asphalt mixture has been developing rapidly all over the world. In some developed countries, especially in Europe and America, many large companies and research institutes have developed their own products [1]. As one of warm mix asphalt technology, the foamed asphalt warm mix technology has been highly praised by experts and scholars all over the world due to the advantages of no additional additive, high economic efficiency and good environmental benefits.

Foamed asphalt is not a new asphalt binder. When asphalt at high temperature and cold water are mixed, it will produce slight bubbles and expand volume. In ordinary asphalt pavement construction, some measures were taken to avoid foaming of asphalt [2]. However, in 1956, Professor Dr. Ladis H. Csanyi discovered the possibility of using foamed asphalt to treat the stabilized soil at the engineering test station of Iowa State University. After the asphalt foaming, the viscosity of the asphalt was greatly reduced, while the adhesion of the asphalt with the cold wet aggregate increased. In

Wen Xiaobo—(1989-), Male, Weinan, Shanxi Province, Master, Majoring in asphalt pavement maintenance and sustainable development technology research work. Fund Project: Jiangsu Natural Science Foundation Project, BK2017015; The Natural Science Foundation of Jiangsu Province (Grants No BK20170156)

1968, Mobil Oil (Australia) improved its original process by replacing steam with cold water and injecting it into hot asphalt under a lower pressure system, making the technology practical and applicable [3].

Now, foamed asphalt can be used for stabilizing various materials, such as crushed aggregate, recycled materials and high plasticity materials. With the wide use of modified foamed asphalt in the highway increasing, how to solve the foaming process, determine the best parameters, and produce asphalt with better performance and higher benefit ratio becomes the key problem that needs to be solved.

In order to make best use of the technology, the researches about the comparison and mechanism analysis of decay functions of two types of imported bitumen was conducted, which found decay curve could be used to characterize the foaming properties of foamed asphalt [4]. Ke [5] and Ali [6] studied the effect of forming parameters on foaming quality and performance of foamed warm mix asphalt respectively by laboratory test. To have better understanding of the technology, the effect of main control parameters on asphalt foaming quality was analyzed by various of means, such as fluent simulation, non-duplication double repetition element variance analysis [4, 7, 8].

Although some studies have been carried out, the foaming mechanism and the optimum foaming parameters need to be further determined. In this paper, the foaming mechanism of modified asphalt was studied by using Wittgen foaming equipment and the foaming process of modified asphalt and factors affecting the quality were analyzed through a large number of experiments and mathematical analysis. Finally, the optimum foaming parameters of SBS modified asphalt were determined.

2 Asphalt Foaming Technique

2.1 Principle of Asphalt Foaming

During the whole process of asphalt foaming, chemical properties of asphalt have not changed, but only a temporary change in physical properties. When the asphalt at high temperature meets cold water, the heat between them will transfer. The temperature of the cold water will rise and the temperature of the asphalt will decrease. When the water is heated to a certain extent, it will become water vapor. The volume of water vapor expands in the asphalt, which will increase the interface of asphalt because of the specific surface area of the asphalt increasing [9]. The asphalt binder wrapped with water vapor is ejected from a specific foaming equipment at a certain pressure to form a vapor-coated asphalt bubble. The surface tension of the asphalt keeps the water vapor in the asphalt and maintains a relatively balanced state, which will last for a few seconds. As time goes on, the asphalt foam will burst, so the foamed asphalt can only exist for a certain period of time. When the foam is in existence, the viscosity is relatively low, so it is suitable for mixing at low temperature. After the temperature of asphalt decreases, the asphalt foam will gradually burst, and eventually the asphalt will become the same binder as ordinary asphalt.

2.2 Evaluation Parameters of Asphalt Foaming

Two indexes have been used to evaluate the foaming effect of foamed asphalt widely, which is half-life and expansion ratio of foamed asphalt. The half-life of foam asphalt refers to the time consumed when the volume of foam asphalt is reduced to half of the volume. The expansion ratio of foamed asphalt refers to the ratio of the maximum volume of foamed asphalt to the primary volume [10]. The two parameters can be measured by a measuring cylinder attached to the equipment stopwatch in laboratory.

The foaming conditions require the large expansion ratio and long half-life to get good quality of foamed asphalt. In this paper, a large number of laboratory tests on foamed asphalt have been carried out in order to obtain the best foaming conditions. The optimum foaming conditions of foamed asphalt needs to meet the requirements of half-life and expansion ratio at the same time. The best foaming conditions of foam asphalt include the optimal water consumption, asphalt temperature and water temperature.

The factors influencing the asphalt foaming quality are as follows:

- (1) Asphalt temperature. Increasing the foaming temperature can improve the flowability of asphalt and provide the heat needed for water vaporization. Increasing the temperature of asphalt is conducive to foaming of asphalt. During the laboratory test, it was found that when the foaming temperature of SBS modified asphalt was lower than 140 °C, it was difficult to foam and easy to block the pipeline of asphalt foaming equipment. However, too high temperature will also cause adverse effects. Therefore, it is necessary to determine the optimum foaming temperature.
- (2) Water consumption. For the foamed asphalt, the most important factor is the water consumption of asphalt. Determining the optimal water consumption of foaming is also the focus of the experiment, because the half-life and expansion ratio of the foamed asphalt are affected by the change of water volume, so it is particularly important to determine the optimal water consumption to prepare the foamed asphalt.
- (3) Water temperature. Water temperature will also have a great impact on the test of foamed asphalt. When hot water is added to the asphalt, the heat transfer will be small, and the instantaneous viscosity will also decrease slowly.

2.3 Foaming Equipment

This paper adopted the WLB10S indoor asphalt foaming test machine produced by Wirtgen company of Germany, which could be used for asphalt foaming test and foamed asphalt mixture production. The injection system was the system used by WR2500, which could simulate the real situation of site construction [11]. The laboratory's asphalt foaming equipment should be able to produce foamed asphalt continuously under certain water content. The quality of the foamed asphalt produced by each batch of asphalt foaming equipment could reach 10–20 kg (Figs. 1 and 2).



Fig. 1. WLB10S asphalt foaming experimental equipment



Fig. 2. Expansion of foamed asphalt

3 Asphalt Foaming Equipment

3.1 Raw Material Properties

Before the foaming test of SBS modified asphalt, the properties of the modified asphalt used in the test were evaluated. The results and requirements of SBS modified asphalt were shown in Table 1. In order to ensure the effect of construction treatment, the softening point, segregation and dynamic viscosity requirements of modified asphalt were improved appropriately.

Table 1. Test results and requirements of modified asphalt

Properties	Unit	Results	Criteria	Test method
Penetration (25 °C, 100 g, 5 s)	0.1 mm	54	50–80	T0604-2011
PI	–	–0.13	–1.2	T0604-2011
Ductility (5 °C, 5 cm/min)	cm	40	≥ 40	T0605-2011
Softening point TR&B	°C	71.5	≥ 70	T0606-2011
Dynamic viscosity (60 °C)	Pa·s	>13000	–	T0625-2011
Rotational viscometer (135 °C)	Pa·s	2.43	≤ 3	T0625-2011
Softening point difference in 48 h	°C	1.2	≤ 2.0	T0661-2011
Elastic recovery (25 °C)	%	99.65	≥ 75	T0662-2000
Solubility	%	98	≥ 99	T0607-2011
Flash point	°C	328	≥ 230	T0611-2011
Mass Loss	%	0.01	≤ ± 1.0	T0610-2011
Penetration ratio (25 °C)	%	81.2	≥ 55	T0604-2011
Ductility (5 °C)	cm	31	≥ 25	T0605-2011

3.2 Orthogonal Test

The foaming process is mainly the change of physical properties of asphalt. At a certain temperature and water consumption, each kind of asphalt will have an optimal foaming state. In the paper, the best foaming effect of modified asphalt was explored based on the foaming effect, and finally the best foaming parameters of modified asphalt were determined.

The preparation of foamed modified asphalt was mainly affected by the following factors:

- (1) Foaming temperature: When SBS modified asphalt foaming test was carried out in laboratory, it was found that it was easy to cause blockage of foaming machine and difficult to foam when the temperature was lower than 140 °C, so the lowest temperature of foaming test of modified asphalt was set at 150 °C. At the same time, considering the aging of asphalt caused by excessive heating temperature, the foaming test temperature of modified asphalt was set at 150–180 °C.
- (2) Water consumption: According to a large number of literature in China and abroad, the results showed that the optimum foaming water consumption of asphalt was about 2.0%, and the optimum foaming water consumption of modified asphalt was slightly higher than that of ordinary asphalt [12]. Therefore, in order to investigate the effect of different water consumption on foaming effect of modified asphalt, four different water consumption were selected, which were 1.5%, 2.5%, 3.5% and 4.5% respectively.
- (3) Water temperature: Considering the foaming principle of asphalt, the heat exchange rate between water and asphalt would be different and affect the foaming effect of modified asphalt. Therefore, the foaming test of SBS modified asphalt was carried out by adding water at 10 °C, 30 °C, 50 °C and 70 °C respectively in Wirtgen foaming equipment. The final parameters and gradient of the test were shown in Table 2.

Table 2. Influencing factors of asphalt foaming test

Influencing factors	SBS modified asphalt
Foaming temperature	150 °C, 160 °C, 170 °C, 180 °C
Water consumption	1.5%, 2.5%, 3.5%, 4.5%
Water temperature	10 °C, 30 °C, 50 °C, 70 °C

From Table 2, it can be seen that the foaming test of modified asphalt was a three-factor and four-level test. If a comprehensive test was carried out to study the interaction of various factors, it would be necessary to conduct $4 \times 4 \times 4 = 64$ experiments, which would inevitably lead to too much work. Orthogonal experimental design can just solve the problems. The idea of orthogonal experimental design is to select some representative points from a comprehensive experiment to carry out the experiment. The principle of selecting representative points is uniform and comparison in orthogonal experimental design. The orthogonal test is adopted for its efficiency, speed and economy, which is mainly designed by fractional factorial design.

According to orthogonal test design, L16 (4^3) orthogonal table was established for the foaming test of modified asphalt. The test was carried out and the results were shown in Table 3.

Table 3. Expansion ratio and half-life of modified foamed asphalt

Group	No.	Water consumption (%)	Asphalt temperature (°C)	Water temperature (°C)	Expansion ratio (times)	Half-life (s)	Average expansion ratio (times)	Average half-life (s)
1	1	1.5	150	10	5.2	35.3	5.5	35.4
	2				5.8	34.8		
	3				5.4	36.2		
2	1	1.5	160	30	6.3	35.2	6.3	34.8
	2				6.8	34.2		
	3				5.9	34.9		
3	1	1.5	170	50	6.4	33.9	6.6	34.2
	2				6.8	34.3		
	3				6.6	34.3		
4	1	1.5	180	70	6.5	30.9	6.5	33.4
	2				6.9	34.5		
	3				6.2	34.7		
5	1	2.5	150	30	9	30	8.9	30.1
	2				9.3	30.5		
	3				8.3	29.7		
6	1	2.5	160	10	10.3	30.8	9.9	30.5
	2				10.1	30.7		
	3				9.4	30		
7	1	2.5	170	70	9.8	28.9	9.6	29.4
	2				9.2	29.7		
	3				9.7	29.7		
8	1	2.5	180	50	9.7	28.8	9.5	29.1
	2				9.3	29.4		
	3				9.4	29.1		
9	1	3.5	150	50	10	30.5	9.5	30.4
	2				9.1	30.5		
	3				9.4	30.3		
10	1	3.5	160	70	10.2	29.8	9.9	29.3
	2				9.7	29		
	3				9.9	29.1		
11	1	3.5	170	10	10.1	29.8	10.5	29.5
	2				11	29.8		
	3				10.5	28.9		
12	1	3.5	180	30	10.1	29.1	10.0	29.1
	2				10.2	29.2		
	3				9.8	29.1		
13	1	4.5	150	70	8.9	28.5	8.7	27.9
	2				7.9	27.3		
	3				8.4	27.9		

(continued)

Table 3. (continued)

Group	No.	Water consumption (%)	Asphalt temperature (°C)	Water temperature (°C)	Expansion ratio (times)	Half-life (s)	Average expansion ratio (times)	Average half-life (s)
14	1	4.5	160	50	8.9	26.9	8.4	27.5
	2				8.2	28.1		
	3				8.1	27.4		
15	1	4.5	170	30	7.8	26.1	8.1	26.1
	2				8.4	26.4		
	3				8.3	25.9		
16	1	4.5	180	10	7.8	25.8	7.7	25.8
	2				8.4	26.3		
	3				8.3	25.2		

The following conclusions could be gained from the test data:

- (1) The foaming temperature of asphalt had no inevitable relationship with the foaming effect. It was not that the higher the foaming temperature of modified asphalt was, the better the foaming effect would be. The increase of foaming temperature of asphalt would have a beneficial effect on the foaming of asphalt, but the viscosity of asphalt would decrease at the same time, which would affect the elasticity of asphalt foam and had a negative impact on the foaming effect.
- (2) For modified asphalt, when the temperature condition was same, with the increase of water consumption, the expansion ratio of modified asphalt increased, while the half-life decreased. This showed that the change of water consumption would have a greater impact on the foaming effect of foamed asphalt, which must be determined to get a suitable expansion ratio and half-life parameters.
- (3) According to the analysis of the factors affecting the foaming characteristics of asphalt, these factors had an interactive effect on the foaming, and the quality of each factor could not be determined solely. Therefore, it was necessary to optimize the evaluation method of foaming effect and analyze the influence of these factors on foaming performance comprehensively.

4 Response Surface Analysis Method

4.1 Analysis Method

As a combination of statistical and mathematical methods, response surface methodology is an optimization method for determining the best level range. A model will be established by multi-factors affecting each other and controlled by continuous variables. Through the method, experimental groups, manpower, material and financial resources can be saved. Furthermore, response surface analysis regards the response of the system as a multi-factor function and shows the function relationship by graphic technology. Therefore, the optimal test conditions can be determined by visual inspection, which has

the advantages of simplicity, rapidity, high efficiency and accuracy. It has been applied in many optimization schemes and succeed.

In the paper, the optimum preparation parameters of foamed asphalt were determined by response surface methodology. The main steps were as follows:

- (1) All the data that met the half-life and maximum expansion ratio requirements were chosen and arranged in Excel tables, then imported into the structure matrix of MATLAB. In the process of data entry, it should be noted that the row matrix should be the same as the original matrix. After data entry, it was necessary to save the data for future use.
- (2) Then the objective function was determined. The maximum expansion ratio and half-life were used as the objective function of the experiment respectively. Among them, Y1 represented the maximum expansion ratio of SBS modified asphalt, and Y2 represented half-life.
- (3) The least square method was used to fit the objective function equation with the orthogonal experimental data, and the variance analysis of the objective function equation was also conducted with MATLAB software to verify the reliability and significance level of the model. If the p-value of variance analysis was less than 5%, the model met the significance level, otherwise it needed to be redesigned until it meets the requirement. The objective function was expressed by second-order polynomial model as follows.

$$Y = a_0 + \sum_{i=1}^n a_i X_i + \sum_{i=1}^n a_{ii} X_i^2 + \sum_{i < j} a_{ij} X_i X_j$$

Where Y is he objective function, n is the number of variables (n = 3), a₀ is a constant, a_{ii} is a quadratic parameter coefficient, a_{ij} is a cross-term coefficient.

- (4) According to the mathematical model established by the above polynomial, the three dimensional spline interpolation was adopted to obtain the continuous 3D surface figure of the model through the Surf function of MATLAB software, and the optimal preparation conditions of foamed SBS modified asphalt could be observed.

4.2 Mathematical Model of Expansion Ratio

According to the test data selected by response surface analysis method, the second-order mathematical model of maximum expansion ratio of SBS modified asphalt was established and the variance analysis of the model was carried out.

$$Y_1 = -74.942 - 1.25X_1^2 - 0.002X_2^2 + 1.56E^{-4}X_3^2 + 13.459X_1 + 0.755X_2 + 0.014X_3 - 0.034X_1X_2 - 3.41E^{-4}X_1X_3 + 0.006X_2X_3$$

(R² = 0.995, p = 3.75E - 06)

The results of variance analysis were shown in Table 4. The table shows that the p-value of ANOVA was 3.75E-06, which met the test requirement of significance level. The residual of ANOVA was plotted in Fig. 3.

Table 4. Variance analysis of maximum expansion ratio of modified foamed asphalt

Type	Correlation coefficient R^2	Means	Mean square deviation	F-value	p-value
Expansion ratio	0.995	8.475	1.5455	127.1938	3.75E-06

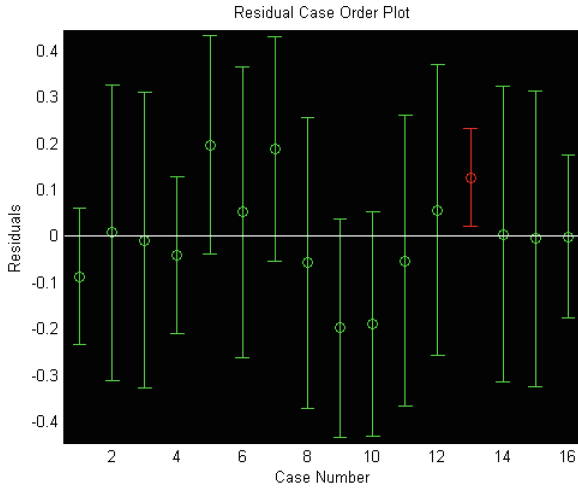


Fig. 3. Residual analysis of expansion ratio of SBS modified foamed asphalt

As shown in Fig. 3, all the points except for the thirteenth (in red) point are near the residual zero, and the confidence intervals for all the points include zero, which shows that the regression model can well fit the expansion ratio of modified asphalt.

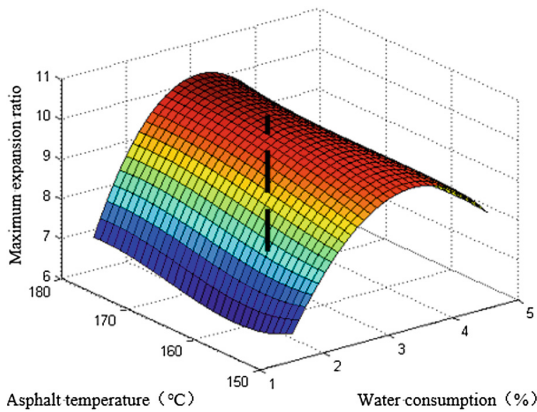


Fig. 4. Interactive surface of expansion ratio of modified foamed asphalt

It can be seen from Fig. 4 that the maximum expansion ratio of modified asphalt increases with the increase of water consumption, and the increasing rate decreases continuously. The temperature of asphalt also has a certain influence on the maximum expansion ratio, but the influence is small. The coordinate position of maximum expansion ratio can be obtained through MATLAB 3D surface elevation map.

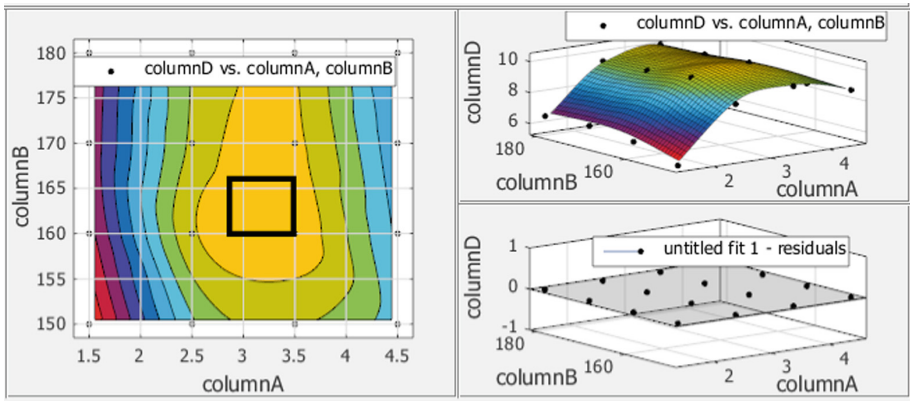


Fig. 5. 3D elevation chart of modified asphalt expansion ratio

According to Fig. 5, the maximum expansion ratio of SBS modified asphalt could be obtained from the surface elevation map. The maximum expansion ratio occurred at the asphalt temperature of 160 °C–165 °C and the water consumption occurs at 3.0–3.5%.

4.3 Mathematical Model of Half-Life

According to the test data selected response surface analysis method, the second-order mathematical model of half-life of modified asphalt was established. The variance analysis of the model was also carried out. The result of variance analysis was shown in Table 5. The p-value of variance analysis was 0.0086, which met the test requirements of significance level, and the variance score was also given. The residuals of the analysis were plotted and illustrated in Fig. 6.

$$Y_2 = 25.561 + 0.481X_1^2 - 6.3E^{-5}X_2^2 - 1.6E^{-5}X_3^2 - 2.506X_1 + 0.120X_2 + 0.246X_3 - 0.021X_1X_2 - 0.002X_1X_3 + 0.011X_2X_3 \quad (R^2 = 0.927, p = 0.0086)$$

Table 5. Variance analysis of half-life of modified foamed asphalt

Type	Correlation coefficient R ²	Means	Mean square deviation	F-value	p-value
Half-life	0.927	30.1536	2.9173	8.4631	0.0086

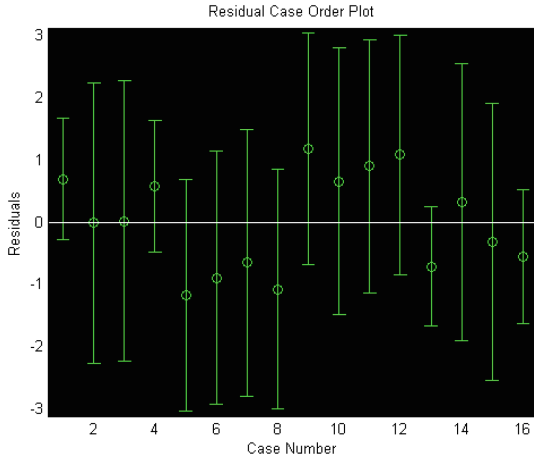


Fig. 6. Residual analysis of half-life of modified foamed asphalt

From the results, it can be seen that there are no abnormal points in the residual diagram of modified asphalt, and all the data are close to the zero, indicating that the regression function of halt half-life can meet its original data well.

According to Fig. 7, it indicates that the half-life of modified asphalt is affected by the temperature of asphalt and the amount of water. The half-life value decreases with the increase of the amount of water consumption. The temperature of asphalt has no significant effect on the half-life.

In summary, the parameters had influence on the half-life of modified foamed asphalt, but the effect was not significant. The half-life values were around 30 s and could meet the requirements of the specifications.

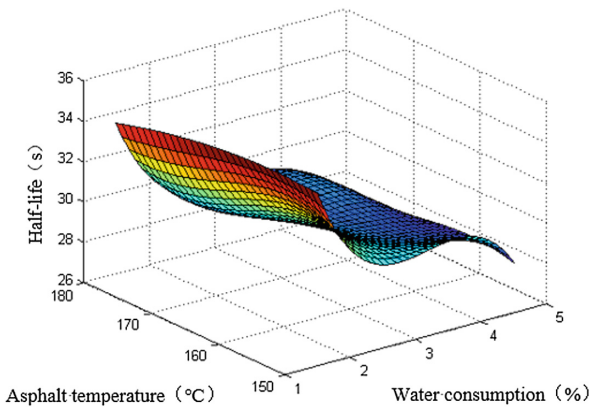


Fig. 7. Interaction surface of foamed half-life of modified asphalt

Therefore, the foaming parameters needed to meet the maximum requirement of expansion ratio first, which could be the best condition for foaming. According to the Figs. 4 and 5, the optimum foaming parameters of modified asphalt were determined as follows: heating temperature, water consumption and water temperature were 165 °C, 3% and 30 °C respectively.

5 Conclusion

Through the study of the foaming characteristics of modified asphalt, the following conclusions were obtained:

- (1) The foaming properties of modified asphalt were evaluated by half-life and expansion ratio. Because of the interaction of different factors, only when the two are relatively large could obtain good foaming effect.
- (2) For modified asphalt, under the same temperature condition, with the increase of water consumption, the expansion ratio of modified asphalt increased, while the half-life decreased. Therefore, once the water consumption of asphalt foaming was changed, the foaming effect of foamed asphalt would also change greatly. The optimum expansion ratio and half-life could be reached only when the suitable water consumption range was determined.
- (3) MATLAB software was used to fit the data of foaming parameters of modified asphalt with second-order mathematical model, and response surface analysis was adopted under the condition of satisfying the requirement of notability. The foaming parameters of modified asphalt were determined as: heating temperature, water consumption and water temperature were 165 °C, 3% and 30 °C respectively.

References

1. Khosravifar, S., Schwartz, C.W., Goulias, D.G.: Mechanistic structural properties of foamed asphalt stabilized base materials. *Int. J. Pavement Eng.* **16**(1), 27–38 (2015)
2. Yuehuan, Y.: *Experimental Study on Warm Mix Asphalt*. Hebei University of Technology (2014)
3. Shu, X., Huang, B., Shrum, E.D., et al.: Laboratory evaluation of moisture susceptibility of foamed warm mix asphalt containing high percentages of RAP. *Constr. Build. Mater.* **35** (10), 125–130 (2012)
4. Cuixing, C., Guiping, H., Xin, Q.: Comparison and mechanism analysis of two kinds of decay equation of imported asphalt. *China J. Highw.* **18**, (3) (2005)
5. Ke, L., Jie, J.S.: Optimization of process parameters for foamed asphalt with different materials prepared by cold recycling equipment. *J. Chang'an Univ. (Nat. Sci. Ed.)* **34**(1), 13–17 (2014)
6. Ali, A., Abbas, A., Nazzal, M., et al.: Effect of temperature reduction, foaming water content, and aggregate moisture content on performance of foamed warm mix asphalt. *Constr. Build. Mater.* **48**(Complete), 1058–1066 (2013)

7. Qichao, W.: Experimental and Simulation Study on Technological Parameters of Asphalt Foaming. Chang'an University (2012)
8. Yi, F., Cheng, Z., Yongyun, C.: Application of grey relational theory in optimization of key parameters of foamed warm mix asphalt production. *J. Guizhou Univ. (Nat. Version)* **31**(2), 119–123 (2014)
9. Feng, Z., Fen, Y.: Foreign warm-mix asphalt mixture technology and performance evaluation. *Chin. Foreign Highw.* **27**(6) (2007)
10. Sunarjono, S.: Performance of foamed asphalt under repeated load axial test. *Procedia Eng.* **54**, 698–710 (2013)
11. Rui, D.: Research on Reliability and Optimization Method of Asphalt Concrete Pavement Structure Based on Numerical Simulation. Chongqing Jiaotong University (2009)
12. Hasan, M.R.M., Shu, W.G., You, Z.: Comparative study on the properties of WMA mixture using foamed admixture and free water system. *Constr. Build. Mater.* **48**(48), 45–50 (2013)



Adaptive Fuzzy Logic Traffic Signal Control Based on Cuckoo Search Algorithm

Suhua Wu¹, Yunrui Bi^{1,2(✉)}, Gang Wang³, Yan Ma¹, Mengdan Lu¹,
and Kui Xu¹

¹ School of Automation, Nanjing Institute of Technology, Nanjing, China
biyunrui@njit.edu.cn

² Key Laboratory of Measurement and Control of CSE, Ministry of Education,
Southeast University, Nanjing, China

³ School of Business, Qingdao University, Qingdao, China

Abstract. Traffic congestion becomes a big problem to perplex the current society. Effective traffic signal control can alleviate traffic congestion, especially for real-time traffic signal control. To improve the control efficiency, fuzzy logic control based on cuckoo search algorithm is applied to solve the problem of real-time traffic signal control. Research object is multi-lane four-phase single intersection which is also the commonly intersection in reality. Vehicular evaluation index model is established firstly. Then, the appropriate green time is given by the cuckoo search algorithm and fuzzy logic control according to the number of real-time road vehicles. Through simulation experiments, the proposed method based on the fuzzy logic control optimized by cuckoo search algorithm can be verified to obtain a good effect. This method also suits for other complex nonlinear systems.

Keywords: Fuzzy logic control · Cuckoo algorithm ·
Real-time traffic signal control

1 Introduction

With the development of the society, traffic jam becomes more and more serious. The effect of traffic signal control is one of the important factors for traffic jam [1]. The development of traffic signal control has experienced three stages, namely fixed-time control, induction control and adaptive traffic control. Fixed-time control method allocates green time according to the historical traffic data distribution. It is suitable for the intersections whose traffic flow change is regular, and the pattern of traffic demand is fixed. However, it will not respond to short-term changes. Induction control expands the green light time by measuring the real-time traffic arrival situation through the vehicle detector. It solves some limitations of fixed-time control to some extent, but this controller only considers extending the time of the current passing phase, while ignoring the waiting vehicles of other phases. In order to overcome this limitation, the adaptive traffic control emerges based on the present and the past traffic information using artificial intelligent control technology [2–9], such as neural networks, fuzzy logic control, various kinds of evolutionary algorithms.

Among of that, fuzzy logic control mimics human thinking and translates the expert knowledge into computable numerical data without the necessity to setup a precise mathematical model or define an exact relationship between input and output variables. It can be observed that a rapidly growing interest using fuzzy control in the field of traffic signal control [7, 10, 11]. However, the set of parameters in fuzzy logic system is extremely important to influence the effect of control. Optimizing parameters using intelligent evolutionary algorithms is one of the important way to determine the values of parameters.

Cuckoo search (CS) algorithm, also called cuckoo search, is a new heuristic algorithm put forward by Cambridge university professor Yang and Deb in 2009 [12]. The idea is based on two main strategies: the cuckoo’s nest parasitic and Levy flight (Levy flights) search mechanism. By means of random walk search for an optimal bird’s nest to incubate their eggs, this approach can achieve an efficient optimization model. The main advantages of CS algorithm are high global search ability, less parameters, better optimization path, high ability of multi-objective optimal problem, good generality and robustness. So, it has attracted many attentions of the scholars [13–15]. In this paper, a kind of fuzzy logic control optimized by CS algorithm is proposed to conduct the traffic signal control problem.

2 Traffic Evaluation Index Model

The most common multiple lanes single intersection is selected as study object, which has universal practical significance for its research. The intersection includes north-south straight traffic flow, north-south right-turn traffic flow, north-south left-turn traffic flow, west-east straight traffic flow, west-east right-turn traffic flow and west-east left-turn traffic flow.

Four-phase traffic signal control is adopted which is also the most common form in reality, as shown in Fig. 1, that are, “west-east straight phase”, “west-east left-turn phase”, “north-south straight phase” and “north-south left-turn phase”.

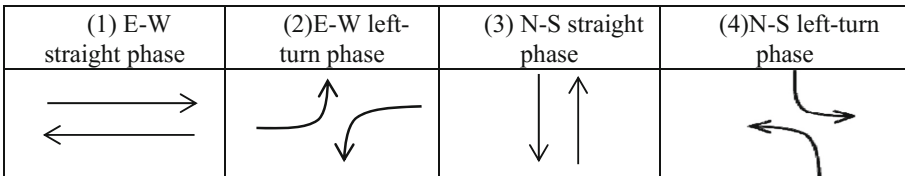


Fig. 1. Phase diagram of a four-phase signal

Suppose that the minimum successive time unit is 1 s, and two vehicles are not allowed to arrive at the same lane in a time unit. Then the number of arrived vehicles during a time unit is expressed:

$$q(i) = \begin{cases} 1, & \text{if a vehicle arrives during the } i\text{th time unit} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

For any moment, four-phase traffic signal control contains one green phase and three red phases. Suppose V_G denotes the number of waiting vehicles at the beginning of the green phase, and V_R^1, V_R^2, V_R^3 stand for the numbers of waiting vehicles at the beginning of the three red phases respectively. So the queue lengths of the green phase and the three red phases in the n th time unit can be depicted as

$$Q_{Gn} = z * \left\{ \left[\left(V_G + \sum_{i=1}^n q_i \right) / p \right] - s * n \right\} \quad (2)$$

$$Q_{Rn}^j = \left[\left(V_R^j + \sum_{i=1}^n q_i \right) / p \right], j = 1, 2, 3 \quad (3)$$

where p is the number of lanes; s is saturation flow rate, and $s = 1$ is supposed here; $z = 1$ if $\left[\left(V_G + \sum_{i=1}^n q_i \right) / p \right] - s * n \geq 0$, otherwise $z = 0$.

Considering the vehicles at the front of line will spend much more time to pass the intersection, so assume that the first car will spend 1.9–2.1 s, the second car will spend 1.75–1.9 s, the third car will spend 1.6–1.75 s, the fourth car will spend 1.45–1.6 s, the fifth car will spend 1.3–1.45 s, the sixth car will spend 1.15–1.3 s, the seventh car will spend 1.05–1.15 s, and the cars after the seventh car will spend 1 s. In order to facilitate the calculation of average vehicle delay, it equivalent to all cars wait for w seconds. If the number of cars in green phase is more than eight, the waiting time w should be $((0.9-1.1)*7 + (0.75-0.9)*6 + (0.6-0.75)*5 + (0.45-0.6)*4 + (0.3-0.45)*3 + (0.15-0.3)*2 + (0.05-0.15))/7 = 2.4-3.1$. If the number of cars m is less than 8 vehicles in green light phase, $w = (m + (0.9-1.1)*(0.75-0.9)*((m-1) + |m-1|)/2 + (0.6-0.75)*((m-2) + |m-2|)/2 + (0.45-0.6)*((m-3) + |m-3|)/2 + (0.3-0.45)*((m-4) + |m-4|)/2 + (0.15-0.3)*((m-5) + |m-5|)/2 + (0.05-0.15) * ((m-6) + |m-6|)/2)/m$, ($0 < m < 8$), if $m = 0$, $w = 0$.

Intersection's vehicle delays can be expressed as the sum of vehicular queue length at each moment. Therefore, the vehicular delays of the current green phase and the three red phases can be expressed respectively as follows

$$D_{Gn} = \sum_{k=1}^n z * \left(V_G + \sum_{i=1}^k q_i - s * k * p \right) \quad (4)$$

$$D_{Rn}^j = \sum_{k=1}^n \left(V_R^j + \sum_{i=1}^k q_i \right), j = 1, 2, 3 \quad (5)$$

In Formula (4), $z = 1$ if $V_G + \sum_{i=1}^k q_i - s * k * p \geq 0$, otherwise $z = 0$. Then the total delay of the current phase is

$$D = D_{Gn} + D_{Rn}^1 + D_{Rn}^2 + D_{Rn}^3 \tag{6}$$

For four-phase intersection, each cycle consists of four phase, so the total vehicular delay of each cycle is the sum of four phases' delay. Therefore, the total delay time during the l th cycle is

$$D^l = \sum_{x=1}^4 D_x \tag{7}$$

where D_x represents the total vehicular delay of the x th phase.

The total vehicular number of the current cycle is the sum of the number of waiting vehicles at the end of the last cycle and the vehicles that reach in this cycle. If q_l indicates the number of vehicles that arrive during the l th cycle, S_l stands for the number of waiting vehicles at the beginning of the l th cycle, then the average vehicular delay of the l th cycle can be expressed as

$$d_l = \frac{D^l}{q_l + S_l} \tag{8}$$

3 Fuzzy Logic Controller for Traffic Signal Control

The structure of fuzzy logic controller is depicted in Fig. 2.

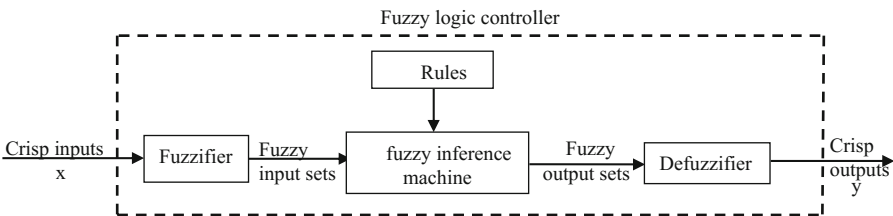


Fig. 2. Structor of fuzzy logic controller

In the process of establishing a fuzzy logic controller, the main works are selecting membership functions and setting rule base. Gaussian membership function's curve shape is smooth. Its control characteristic is more gently. Studies have shown that it has good stability and it is a reasonable form to describe the fuzzy subset. Thus it is adopted in this paper, whose mathematical expression is:

$$\mu_{\bar{A}}(x) = \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad (9)$$

where m presents the center of the membership function, σ is the deviation of the membership function.

The membership function adopts three linguistic partitions, namely “short (S)”, “medium (M)”, “long (L)”. For the input variable QG (QR), they stand for the short, medium and long queue length respectively, while they represent the short, medium and long green light time separately for the output variable T. let the universe of discourse be $[0, 12]$. According to the practical experience, the basic domain of QG (QR) is $[0, 40]$, and the basic domain of T is $[15, 65]$ for the straight phase and $[15, 45]$ for the left-turn phase respectively.

The preliminary rule base is given in Table 1 on the basis of daily experience and the expert knowledge of the traffic police.

Table 1. Basic control level fuzzy rule base

T \ QG	S	M	L
QR	S	M	L
M	S	M	L
L	S	S	M

4 Optimization Process with CS Algorithm

4.1 Principle of CS Algorithm

In order to simulate the habit of cuckoo, Prof. Xinshe Yang et al. assume the following three ideal states of CS algorithm:

- (1) Each cuckoo lays only one egg at a time and randomly selects a nest to hatch it;
- (2) In a randomly selected group of bird nests, the best one will be reserved for the next generation;
- (3) The number of available as the bird’s nest is fixed, the host of a bird’s nest can find the probability of an exotic birds’ eggs $P_a \in [0, 1]$.

By assuming the above three ideal states, the updating formula of the location and path of cuckoo optimization search are as follows:

$$x_i^{(t+1)} = x_i^t + a \oplus L(\lambda), \quad i = 1, 2, \dots, n \quad (10)$$

Where x_t^i refers to the position of the bird's nest in the t th generation of the i th bird nest, \otimes pointing to the point multiplication, a refers to the step length control, which is used to control the step length search range, and its value obeys the normal distribution. $L(\lambda)$ is a random search path for Levy, and the random step is the Levy distribution.

$$L(s, \lambda) \sim s^{-\lambda}, (1 < \lambda \leq 3) \quad (11)$$

where s is the random step length obtained by levy flight.

It can be seen from Formula (10) that the walking mode of the line is a process of random walk. Because of the random walk characteristics of levy's flight, new solutions often appear near the local extreme points, so the short step search of levy's flight is more conducive to improving the quality of the solutions. In addition, there are new solutions at a distance from the local optimal value. Occasionally, large step length exploration makes the algorithm not easy to fall into the local extreme point.

4.2 Controller's Parameters Optimization Based on CS Algorithm

In order to obtain the optimal parameter settings, the fitness function should be set up according to the characteristics of the controlled object. The optimization problem can be converted to solve multi-dimensional function optimal value problem. For traffic control problem, the vehicular average delay at intersection is often considered as an important index for traffic control effectiveness evaluation. According to the established vehicular average delay in Formula (8), the fitness function can be defined as:

$$fitness = \min\left(\left(\sum_{l=1}^M d_l\right)/M\right) \quad (12)$$

where M is the total number of cycles during the simulation time.

In order to reflect the real-time change of traffic flow in time, on-line optimization and adjustment is necessary, which is of great significance for the practical application of control method. We use a signal cycle as an optimization time period. The parameters after optimization are used in fuzzy controller to control the follow-up vehicles in the following time period. The control process is carried out in turn like this.

Summarize the above description, the procedure to optimize the fuzzy logic controller using CS algorithm is addressed as follows:

Step 1: Determine two input variables of the fuzzy logic controller;

Step 2: Initialize encoding for the corresponding membership functions based on the range of each parameter;

Step 3: Calculate the fitness function value of each individual;

Step 4: Carry out the corresponding operations of CS algorithm and produce new population;

Step 5: Judge whether meet the termination conditions, if satisfied, then stop the iteration algorithm. The corresponding parameters are applied to the fuzzy controller to control the vehicles in the following time period; otherwise, return to Step 3.

5 Simulation

5.1 Simulation Conditions

In order to verify the efficiency of the proposed fuzzy control method, it is compared with fixed-time control (FTC) and the traditional fuzzy logic control (FLC). The vehicular average delay and queue length under different methods are compared. According to different levels of traffic congestion, simulation experiments are set under the following six different arrival rates:

Case 1, the arrival rate of straight vehicles is 0.1 veh/s, and the arrival rate of left-turn vehicles is 0.1 veh/s.

Case 2, the arrival rate of straight vehicles is 0.2 veh/s, and the arrival rate of left-turn vehicles is 0.1 veh/s.

Case 3, the arrival rate of straight vehicles is 0.3 veh/s, and the arrival rate of left-turn vehicles is 0.1 veh/s.

Case 4, the arrival rate of straight vehicles is 0.4 veh/s, and the arrival rate of left-turn vehicles is 0.1 veh/s.

Case 5, the arrival rate of straight vehicles is 0.5 veh/s, and the arrival rate of left-turn vehicles is 0.2 veh/s.

5.2 Simulation Results and Analysis

In the above several simulation cases, the comparison results of vehicular average delay under different methods are shown in Table 2. The trend comparisons of vehicular average delay and average queue length under the fixed-time control, traditional fuzzy logic control and fuzzy control optimized by CS algorithm are shown in Figs. 3, 4, 5, 6 and 7.

Table 2. Comparison of vehicular average delay

Arrival rate of the straight phase (veh/s)	Arrival rate of the left-turn phase (veh/s)	Vehicular average delay (s/veh)			Improvement compared with FTC
		FTC	FLC	CSFLC	
0.1	0.1	40.47	33.86	23.79	41.22%
0.2	0.1	43.67	40.95	34.01	22.12%
0.3	0.1	56.58	84.83	42.16	25.48%
0.4	0.1	146.30	143.85	115.11	21.32%
0.5	0.2	238.47	228.03	217.04	8.99%

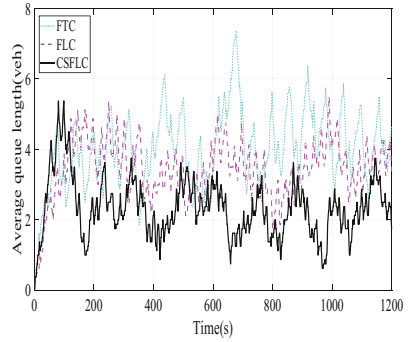
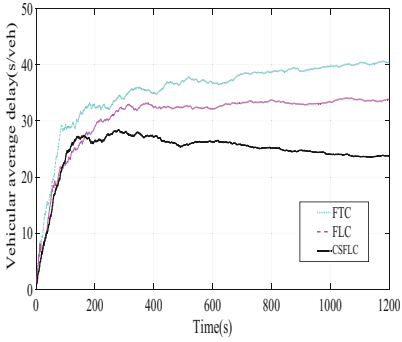


Fig. 3. Vehicular average delay and average queue length comparison under Case 1

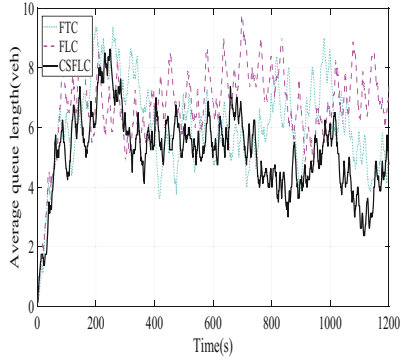
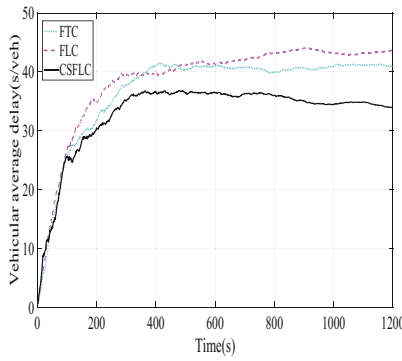


Fig. 4. Vehicular average delay and average queue length comparison under Case 2

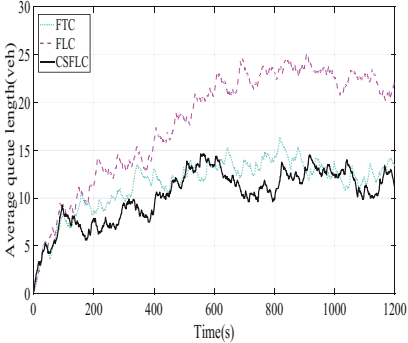
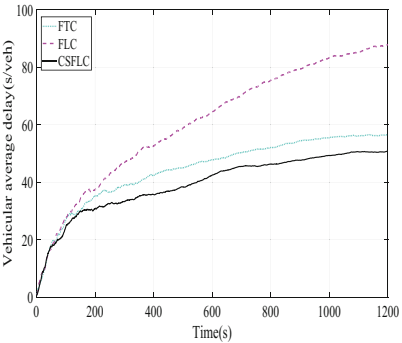


Fig. 5. Vehicular average delay and average queue length comparison under Case 3

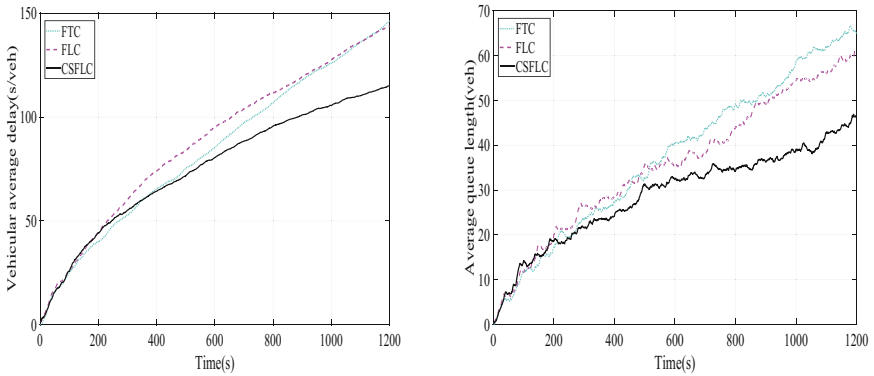


Fig. 6. Vehicular average delay and average queue length comparison under Case 4

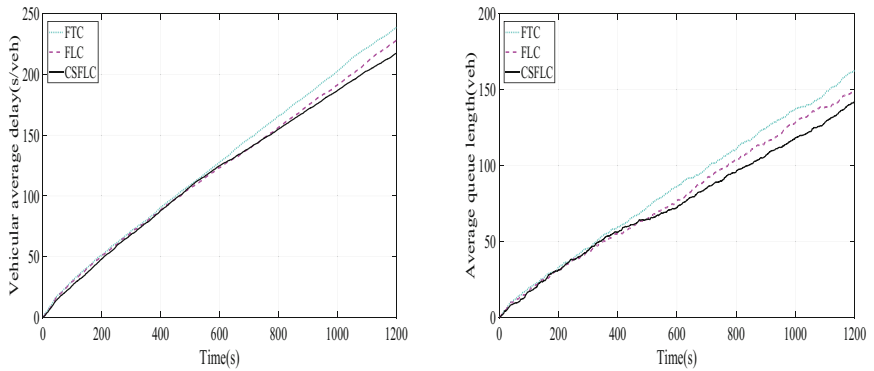


Fig. 7. Vehicular average delay and average queue length comparison under Case 5

From Table 2 and Figs. 3, 4, 5, 6 and 7, we can see that the vehicular average delays and average queue lengths under different methods also increase as the vehicular arrival rate increases. By comparing the vehicular average delays under the same arrival rate among these methods, we can see that fuzzy control method is usually better than fixed-time control method. But when the arrival rate is 0.3, the effect of fuzzy control method is worse than fixed-time control method, which illustrates the importance of parameters and the necessity of parameters optimization. However, the effect of the proposed CSFLC method is the best. When vehicular arrival rate is low (Case1–Case 2), and under slightly crowded state (Case 3–Case 4), the improvement of CSFLC can reach about 21%–41% compared with the fixed-time control. In traffic extremely congestion state (Case 5), vehicles have been far beyond the processing capacity of intersection, so if only by improving the control method at this time, the space to improve is not large, only at 8.99%. But even in this kind of condition, the proposed CSFLC method is also the most effective compared with the other two methods.

6 Conclusion

In order to effectively manage the signal control of the intersections, aiming at the common multi-lane four-phase intersection, this paper established the corresponding traffic evaluation index model, and designed a fuzzy logic controller based on the cuckoo search algorithm, and further apply the optimized fuzzy logic controller to simulate single intersection signal control. The CS algorithm possesses high global search ability, less parameters, better optimization path. It presents better searching capability and maintains diversity of the population. Compared with the fixed-time controller and the traditional fuzzy logic controller, the proposed fuzzy logic controller shows better performance. It can effectively reduce the average vehicle delay time, and alleviate traffic congestion.

It is worth to mention that the proposed method also suits for some other complex nonlinear systems, such as robot control, power systems, servo motor control, etc.

Acknowledgments. The project is supported by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No. 18KJB510016), School-level Project of Nanjing Institute of Technology (No. YKJ201718, No. JCYJ201819), Practical Innovation Training Program of Jiangsu Province for College Students (No. 201811276051X) and a Project Funded by the Priority Academic Development of Jiangsu Higher Education Institutions.

References

1. Srinivasan, D., Choy, M.C., Cheu, R.L.: Neural networks for real-time traffic signal control. *IEEE Trans. Intell. Transp. Syst.* **7**(3), 261–272 (2006)
2. Wang, F.Y.: Parallel control and management for intelligent transportation systems: concepts, architectures and applications. *IEEE Trans. Intell. Transp. Syst.* **11**(3), 630–638 (2010)
3. Balaji, P.G., Srinivasan, D.: Distributed geometric fuzzy multiagent urban traffic signal control. *IEEE Trans. Intell. Transp. Syst.* **11**(3), 714–727 (2010)
4. Zhu, F., Li, Z., Chen, S., Xiong, G.: Parallel transportation management and control system and its applications in building smart cities. *IEEE Trans. Intell. Transp. Syst.* **17**(6), 1576–1585 (2016)
5. Talab, H.S., Mohammadkhani, H.: Design optimization traffic light timing using the fuzzy logic at a Diphasic's Isolated intersection. *J. Intell. Fuzzy Syst.* **27**(4), 1609–1620 (2014)
6. Ding, N., He, Q., Wu, C., Fetzer, J.: Modeling traffic control agency decision behavior for multimodal manual signal control under event occurrences. *IEEE Trans. Intell. Transp. Syst.* **16**(5), 2467–2478 (2015)
7. Benhamza, K., Seridi, H.: Adaptive traffic signal control in multiple intersections network. *J. Intell. Fuzzy Syst.* **28**(6), 2557–2567 (2015)
8. Zhao, D., Dai, Y., Zhang, Z.: Computational intelligence in urban traffic signal control: a survey. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **42**(4), 485–494 (2012)
9. Zhao, Y., Gao, H., Wang, S., Wang, F.: A novel approach for traffic signal control: a recommendation perspective. *IEEE Intell. Transp. Syst. Mag.* **9**(3), 127–135 (2017)
10. Pappis, C.P., Mamdani, E.H.: A fuzzy logic controller for a traffic junction. *IEEE Trans. Syst. Man Cybern.* **7**(10), 707–717 (1977)

11. Murat, Y.S., Gedizlioglu, E.: A fuzzy logic multi-phased signal control model for isolated junctions. *Transp. Res. Part C Emerg. Technol.* **13**(1), 19–36 (2005)
12. Yang, X.S., Deb, S.: Cuckoo search via Lévy flights. In: *Proceedings of World Congress on Nature & Biologically Inspired Computing Coimbatore*, pp. 210–214. IEEE (2009)
13. Rajabioun, R.: Cuckoo optimization algorithm. *Appl. Soft Comput.* **11**(8), 5508–5518 (2011)
14. Liu, X.Y., Fu, M.L.: Cuckoo search algorithm based on frog leaping local search and chaos theory. *Appl. Math. Comput.* **266**, 1083–1092 (2015)
15. Zhu, X., Wang, N.: Cuckoo search algorithm with membrane communication mechanism for modeling overhead crane systems using RBF neural networks. *Appl. Soft Comput.* **56**, 458–471 (2017)



Research on Road Traffic Signal Timing Method Based on Picture Self-learning

Guohua Zhu¹ and Chi Zhang^{2(✉)}

¹ Ze Yi Traffic Engineering Consulting (Shanghai) Co., Ltd., Shanghai, China

² Yan Yun Network Technology (Shanghai) Co., Ltd., Shanghai, China
zhch0507@163.com

Abstract. Based on the pictures of queuing vehicles at intersections collected by video cameras, a method of traffic signal control based on picture self-learning is proposed. Based on Convolution Neural Network, this paper classifies the queuing length pictures of vehicles with different phase-critical traffic flow, establishes the relationship database between the picture data set and the green light display time of the pictures with different phase-critical traffic flow queuing length categories, and obtains the current phase green light display time of the current cycle on the basis of the relationship database, so as to achieve real-time optimization. The purpose of the signal control scheme. This method does not need the exact traffic flow collected by traffic flow detector, but acquires the pictures of queuing vehicles in different periods and phases, and trains the green light duration to control the traffic in real time.

Keywords: Intelligent Transportation Systems · Traffic signal control · Picture self-learning · Convolution Neural Network

1 Introduction

Delays at intersections are the main form of vehicle delays on urban roads. According to statistics, 80% of delays on urban roads are caused by platform intersections. The unreasonable traffic signal control scheme is also an important factor leading to vehicle delay at the platform intersection where traffic signal control is implemented. Therefore, the study of traffic signal control methods has always been the focus of the research in the field of intelligent transportation.

At present, the timing methods of signal control in China are HCM method in the United States, TRRL method in Britain (also known as Webster method), ARRB method in Australia, parking line method recommended in China's "Code for design of urban roads" and so on. Most of these signal timing methods take the maximum traffic rate or the minimum delay as the optimization objective, calculate the signal cycle, and allocate the green light release time of each phase according to the traffic flow. Traffic parameters of traffic signal control system are mainly collected by loop coil, microwave detector and so on. With the maturity of video processing technology and the popularity of video in intelligent transportation, more and more signal manufacturers and Internet enterprises begin to use video as the data source of signal control system. For example, Hicon system of Qingdao Haixin, Smart UTC system of Tianjin Tongxiang,

Intelligent system of Zhejiang University Central Control, Micro system of Shanghai Baokang, City brain project of Alibaba in Hangzhou, "Traffic Go" scheme of Huawei traffic signal optimization, etc.

Video is used as the data source of the signal control system, and the optimization of traffic signal achieves better results. However, there are many shortcomings, such as the algorithm is not open, the video processing is complex, and a large number of video transmission needs high bandwidth network problems, so the investment cost is large. In the actual application of signal control, the signal control system does not require very precise traffic flow. By intercepting pictures with key information from the video, computer vision picture self-learning method can be used to judge the traffic condition at intersections, and then optimize the signal control scheme. In recent years, Convolution Neural Network has achieved remarkable results in computer vision. Krizhevsky team reduced the error record of image recognition classification from 26% to 15% by using convolutional neural network in the ImageNet challenge, greatly improving the classification level of images in 2012 [1]. Wang use machine vision, adopt four STM32F103 as embedded controllers, use cameras to collect traffic flow at intersections, transmit traffic information in four different directions to the controller, the controller collects and makes decisions, realize intelligent control of signal lamp conversion time [2]. Professor Tang of Zhejiang University of Technology has disclosed a traffic signal control method based on deep Convolution Neural Network. By detecting the flow of each exit, the problem of traffic flow spillover at intersection exit is solved [3].

Based on Convolution Neural Network, this paper classifies the vehicle queue length pictures with different phase key traffic flow, and establishes the picture data set and the relation database between the vehicle queue length pictures with different phase key traffic flow and the green light display time. Based on the relational library, the current phase green light display time of the current cycle is obtained to optimize the signal control scheme in real time.

2 Overview of Road Traffic Signal Timing Method Based on Picture Self-learning

Traffic managers often adjust the vehicle passing time at a certain phase of the intersection according to the queuing length of each inlet, optimize the traffic flow in all directions, so as to improve the efficiency of the intersection and reduce the "empty time". This signal optimization method is simple and effective, but it needs the participation of people, and it is difficult to achieve the optimization of 7 * 24 h. In this paper, the images of queuing vehicles are extracted from the video captured by the cameras in the entrance lanes. The Convolution Neural Network is used to classify the queuing states of vehicles in the pictures, and the corresponding relationship between different categories and green light release time is established. The signals are optimized according to the corresponding green light display time of different categories.

Picture-based traffic signal control method assumes that the maximum queue length is at the beginning of the green light. The realization of this control method includes four modules: camera, image processing device, traffic signal controller and traffic

signal lamp. The camera is connected with the picture processing module, the picture processing module is connected with the signal controller, and the signal controller is connected with the traffic signal lamp (Fig. 1).

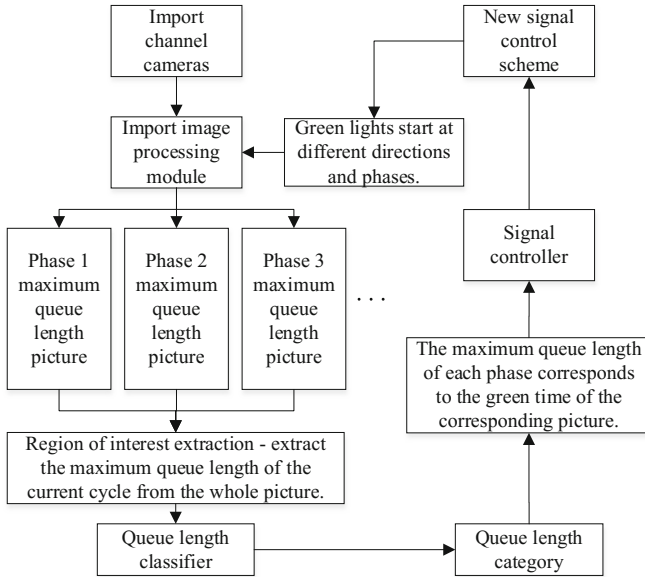


Fig. 1. Composition of traffic control methods.

The pictures collected by the camera are transmitted to the picture processing module, which is connected with the traffic signal controller to obtain the pictures with the maximum queue length of the current signal cycle. Then, the vehicle queue length picture of the current phase corresponding lane is extracted. Finally, the number of queuing vehicles is identified by the vehicle queuing state classifier (the vehicle queuing state classifier is trained by neural network), and the green light display time corresponding to each phase of the vehicle queuing state picture is obtained, which is sent to the signal controller for real-time control of the green light display time.

The main steps of picture-based road traffic signal control method include data set establishment, training, picture classification, and the establishment of the relationship database between picture and green light display time (Fig. 2). The method uses an infrared camera, and the camera automatically switches to the infrared working mode in the case of bad weather (storms, haze) and dark nights. The cameras in all directions of the intersection are extracted and analyzed in the same signal cycle, and the environment for intercepting pictures is the same, and the pictures in different environments can be trained. The method assumes that the number of phases is constant, and only optimizes the phase duration. Considering the adaptability of the long-term control, it is necessary to re-establish a relational library between the classified picture and the green light display time.

In the following four aspects, the acquisition of pictures, the classification of pictures, the relationship between classified pictures and green light display time library, and the determination of green light display time are studied respectively. The signal control method of a cross intersection with special left turn phase is also analyzed.

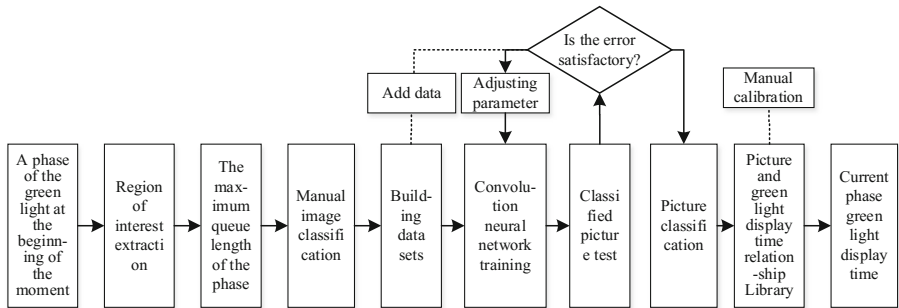


Fig. 2. Traffic signal control method steps.

3 Research on Road Traffic Signal Timing Method Based on Picture Self-learning

3.1 Maximum Queue Length Image Acquisition

In order to get pictures of the maximum queue length of each phase, the cameras need to be installed in suitable locations. The ideal location is in front of the parking line at each entrance, and the camera can take pictures of the entire queue length. In view of the current road intersection electronic police, entrance system installed in the vehicle stop line behind the entrance (usually 18–30 m behind the parking line). The camera can also be mounted on the elevator of the electronic policeman/clip by means of a common rod with the electronic policeman/clip. The principle of camera installation is to see the position of queued vehicle tail. At the beginning of the green light at the end of the red light, the signal controller sends instructions to the picture processing module to extract the current phase of the picture at the current moment, and then normalize the image.

The Region of Interest (ROI) is needed to extract the maximum queuing length of vehicles in each phase. ROI is the area occupied by queuing vehicles at the same phase intersection. ROI region selection, the use of coordinates (select the region of interest of the four points constitute a closed quadrilateral).

3.2 Picture Classification Based on Convolution Neural Network

After the region of interest is extracted, a Convolution Neural Network is constructed to train the classified images, and a vehicle queuing state classifier is obtained, which is used to classify the images with the maximum queuing length. Vehicle queuing state classification based on Convolution Neural Network has stronger anti-interference

ability than traditional technology. It can effectively break through the difficulties of traditional technology, such as sudden change of illumination, complex background, partial occlusion of large vehicles, single scene and so on.

In this paper, the Convolution Neural Network is proposed to detect the vehicle state at intersections, which can avoid the network cost caused by large traffic video transmission and improve the adaptability of vehicle state detection at intersections. Convolution Neural Network consists of input layer, convolution layer, down sampling layer (pool layer), full connection layer and output layer. Sermanet et al. [4, 5] have made some improvements to the basic model of convolutional neural network. By adding the first layer sampling results directly into the input of classifier, more detailed features can be obtained. In this paper, based on the improved network of Sermanet P et al., the queuing state of vehicles at intersections is divided into five categories: VL, L, M, H and VH, which correspond to five nodes in the output layer of the following graph (Fig. 3).

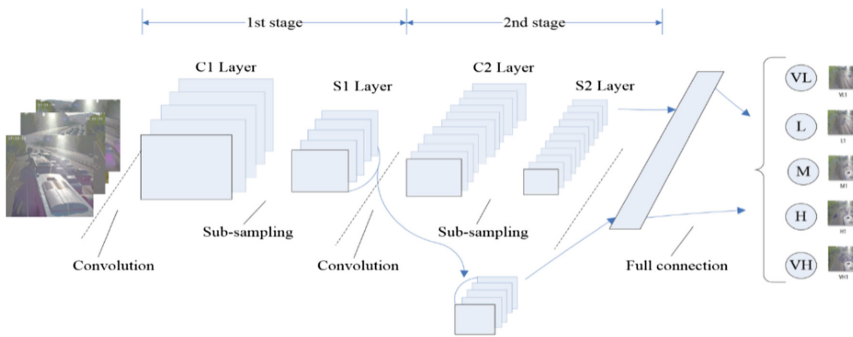


Fig. 3. Multilevel Convolution Neural Network structure for vehicle queue state detection.

3.3 Establishment of Relationship Library Between Classification Images and Green Light Display Time

Based on manual survey and statistics or the theory of vehicle queuing at intersections, the relationship database between classified pictures and green light display time is established, that is, different queue lengths correspond to different green light display time.

At the intersection of vehicle queuing theory as an example, the intersection of the vehicle queue, when the green light vehicle operating characteristics with single lane intersection queue operation characteristics (queuing vehicle headway Fig. 4). When the traffic lights turn green, the vehicles waiting behind the parking line begin to move through the parking line in turn, and the flow rate increases gradually from 0 to saturated flow rate. Thereafter, subsequent vehicles crossing the parking line will remain equal to the saturated flow rate until all vehicles accumulated behind the parking line have been released, or have not been released but the green light time has expired. Based on the headway model of queuing vehicles at one-lane intersections, different green light display times can be obtained according to different queuing lengths.

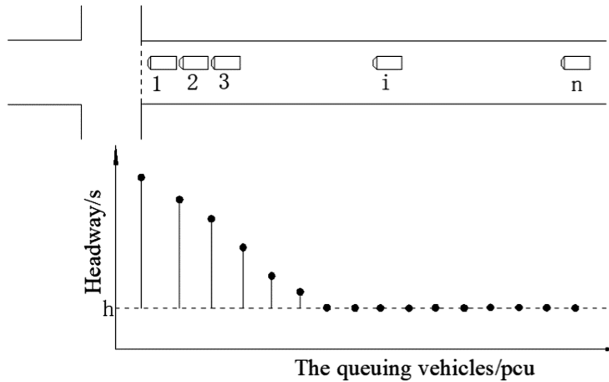


Fig. 4. “Time distance” diagram of queuing vehicle at single lane intersection.

The headway model of queuing vehicles at one-lane intersections is converted to standard cars for study. The length of different types of vehicles converted to Passenger Car Unit (PCU) is approximately equal to the length of vehicles (Table 1). Therefore, the time of vehicles passing through intersections can be studied through different overall queuing states of vehicles.

Table 1. Equivalent car conversion coefficient.

Vehicle type	Conversion coefficient
A passenger car or a freight car less than 3t	1.0
Wagon	1.2
Buses or trucks less than 9t	2.0
9–15t truck	3.0
Articulated bus or large flat trailer truck	4.0

According to the classified pictures, the same class of pictures corresponds to a green light display time, and the queuing state of vehicles is divided into five categories. Then each phase corresponds to five green light display times. Taking the four-phase crossroads with special left-turn as an example, there are 54 (625) different schemes. The more classification of vehicles queuing, the more schemes, the stronger the adaptability of signal control.

3.4 Determination of the Duration of Signal Controlled Green Light at the Current Phase

When the whole red is not considered, the sum of the green light and the yellow light time of each phase is the signal control cycle. For the green light display time of each phase, the minimum green light time such as vehicle passing time, pedestrian crossing time and the maximum green light time of vehicle and pedestrian waiting tolerance should be considered comprehensively. The green light duration and signal control cycle at each intersection are shown in Formula 1.

$$\begin{cases} C = g_1 + g_2 \dots + g_i \dots + g_N + 3N \\ g_{min} \leq g_i \leq g_{max}, \text{ among } i = 1, 2, \dots, N \end{cases} \quad (1)$$

Among them:

- C —the signal control cycle;
- $3N$ —a yellow time is 3 s, N is the number of phases;
- g_i —a green light display time of phase i ;
- g_{min} —the minimum green time;
- g_{max} —the maximum green time.

3.5 Analysis of Signal Control Method for Cross Intersection with Left Turn Special Phase

Through OpenITS [6], an Intelligent Transportation Systems data open website, the video traffic flow detection data of a cruciform intersection with a left-turn dedicated phase is provided, and the signal control method based on picture self-learning is analyzed.

The location of channelization and camera installation at this intersection is shown in Fig. 5. There is one left turn, two straight lanes and one right turn lane at the East entrance, one left turn, three straight lanes and one right turn lane at the West entrance, and two right and straight left lanes at the north and South directions. The signal phase is divided into three phases: East-West, east-west, left and right (east-west direction

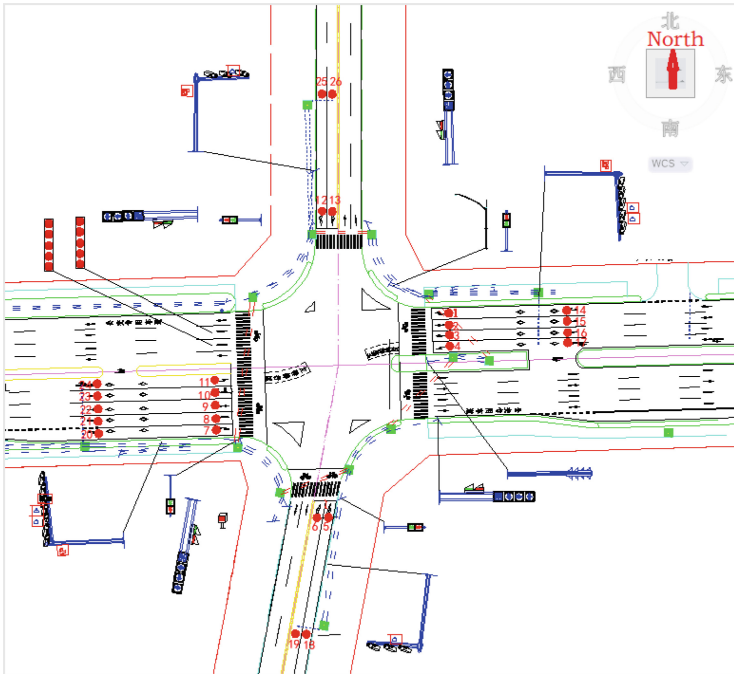


Fig. 5. Channelization of intersection and location of camera installation

with a special left turn phase, and set the left turn waiting area). The camera is installed on 4 inlet roads.

Taking the East-West direct phase as an example, two pictures of the East-West direct direction (west-import direct and East-Import direct) were extracted at the beginning of the end of the East-West direct phase red light, and the pictures were classified. The picture with more queuing vehicles was selected as the key traffic flow, and the pictures were processed by the vehicle queuing state classifier. Classification, through the vehicle queue length category picture and green light display time database to determine the green light display time, other phase of the green light display time determination and method is the same, through this method real-time control green light display time, to achieve the purpose of optimizing signal control scheme. Table 2 shows the straight-line phase classification picture and the green light display schedule of the things obtained by video processing analysis and manual calibration.

Table 2. Cross-shaped intersection, straight line phase classification picture and green light display timetable.

Picture classification	Green light display time/second
VL	12
L	20
M	28
H	36
VH	45

In the straight line phase of the east and west, the green light display time of the VL minimum level classification picture is generally the minimum green time for pedestrians crossing the street, the length of the sidewalk in the east-west direction (crossing the street zebra crossing) is 14 m, and the pedestrian walking speed is generally 1.2 m/s. Therefore, the green light corresponding to VL is displayed for 12 s (14/1.2 rounding). The green light display time of the VH maximum grade classification picture requires the length of the camera to be photographed as far as possible and the pedestrians and drivers waiting in line to determine the endurance time. The camera used at the intersection can shoot up to 100 m, and the green light display time corresponding to VH is 45.

In this intersection, the pictures corresponding to the straight line of a certain period are classified into H level (36 s), the M level corresponding to the left turn (26 s), and the picture corresponding to the north and south straight are classified into L level (18 s). In the case where full red is not considered, according to Formula (1), the signal period duration is 89 s.

4 Summary

This paper proposes a road traffic signal control method based on picture self-learning. The pictures of queuing length of vehicles with different phases are classified by Convolution Neural Network. The picture data set and the relation database between the pictures of queuing length and green light display time of vehicles with different phases are established. On this basis, the current phase green light duration of the current cycle is obtained to achieve the purpose of real-time optimization of signal control scheme, combined with a special left-turn phase of the intersection signal control method analysis.

Road traffic signal control method based on picture self-learning does not require accurate traffic flow detection by traffic flow detector. The green light display time is obtained by obtaining the pictures of queuing vehicles with different phases in different periods, and the real-time traffic control is carried out.

Acknowledgments. The project is supported by Zhangjiang National Independent Innovation Demonstration Zone Special Development Foundation (Number 201705-JD-C1085-072).

References

1. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems, pp. 1097–1105. Curran Associates Inc. (2012)
2. Wang, X., Gao, Y., Liu, D.: Design of intelligent traffic light control system based on machine vision. *Autom. Appl.* **11**, 38–39 (2015)
3. Tang, Y.: A traffic signal control method based on deep Convolution Neural Network. China Patent: CN106023605A, 12 October 2016
4. Sermanet, P., Chintala, S., LeCun, Y.: Convolutional Neural Networks applied to house numbers digit classification. In: 2012 21st International Conference on Pattern Recognition (ICPR), pp. 3288–3291. IEEE (2012)
5. Sermanet, P., LeCun, Y.: Traffic sign recognition with multi-scale convolutional networks. In: The 2011 International Joint Conference on Neural Networks (IJCNN), pp. 2809–2813. IEEE (2011)
6. OpenITS (EB/OL) (2018). <http://www.openits.cn/>



Review of Ramp Metering Methodologies for Urban Expressway

Songxue Gai^(✉), Xiaoqing Zeng^(✉), Chaoyang Wu, and Jifei Zhan

Key Laboratory of Road and Traffic Engineering of the Ministry of Education,
Tongji University, Shanghai 200092, China
{gaisongxue, zengxq}@tongji.edu.cn

Abstract. Ramp metering is a viable, reliable and effective strategy which maintains urban expressways in optimum status. Various forms of applications and researches on ramp metering are studied systematically by literature analysis of various countries. It can be classified into different categories according to different control objects, tasks and methods. In view of the history of ramp metering development, a review and analysis is made for the essential urban expressway ramp metering methodologies and their advantages and disadvantages in this paper, which is based on two aspects: isolated ramp metering and coordinated ramp metering. And also some perspectives on the development of ramp metering research are brought up in order to propose the tendency of further studies.

Keywords: Urban expressway · Ramp metering · Traffic optimization

1 Introduction

The main goal of expressway management is to increase throughput, speed, and capacity to maintain the optimum operation of expressway system. Ramp metering is most commonly expressway management strategy to control the number of vehicles entering an expressway by using traffic signal [1]. Since 1960s, various forms of researches on ramp metering have been done from fixed-time strategy to adaptive or responsive ramp metering. And lots of researches were also put into practice, and also great achievements have been made, especially traffic responsive control [2].

In order to review these essential researches, this paper summarizes the ramp metering strategies and algorithms through three main sections. Firstly, this paper presents the ramp categories according to different control objects, means, and methods; secondly, presents a review of isolated ramp metering strategies and their evaluations which mainly developed in the earlier stage, thirdly provides a review of recent researches in coordinated ramp metering strategies, lastly, gives the conclusion and outlook.

2 Categories of Ramp Metering

The state of the ramp can be generally divided into three types: no control (ramp open), ramp metering (adjustment), and ramp closed. These three states can be converted to one another. Referring to ramp metering, it can be divided into several categories,

according to the control ideas, objects, means and techniques of implementation. As for the specific method to be adopted, it should be determined in accordance with the specific tasks and the actual conditions.

The existing ramp metering strategies and methods are summarized as follows:

(1) According to different objects, it can be divided into isolated ramp metering and coordinated ramp metering; (2) According to different control means, it can be divided into fixed-time control, responsive control and adaptive control; (3) According to different control methods, it can be divided into single control and integrated control.

The above-mentioned strategies and methods can produce multiple combinations to form a ramp metering strategy when in practical use, as shown in Fig. 1. For example, it can combine fixed-time control with responsive control, that some of the multiple channels can be controlled by fixed-time strategy, and others can be controlled by responsive strategy.

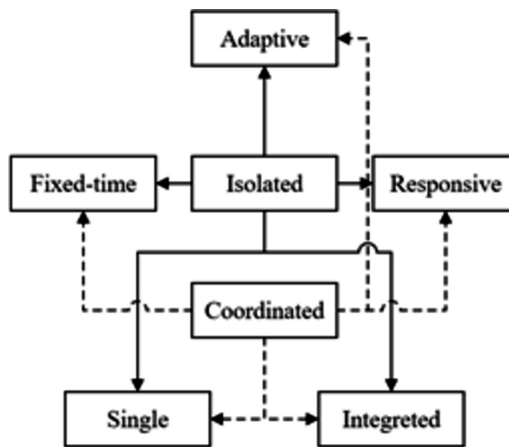


Fig. 1. Classification and combination strategies of ramp metering.

The essential ramp metering strategies and their relationships are shown in Table 1.

Table 1. Ramp metering strategies and representative algorithms.

Object	Strategy	Representative Algorithms
Isolated	Fixed-time strategies	
Isolated	Responsive strategies	D-C\Occupancy\G-A
Isolated	Adaptive control strategy	NN\FL\ADP\GA
Isolated	Integrated optimization Control	Lane-based optimization
Coordinated	With surface road control systems	On-ramp\off-ramp
Coordinated	With other control systems	MTFC\R-G\Monetary Cost
Coordinated	Multi ramp joint control strategy	Bottleneck\LP(NLP)\H\MET
Coordinated	Network flow coordinated control	

3 Isolated Ramp Metering Strategies and Algorithms

Isolated ramp metering (or local ramp metering) is a relatively simple and fundamental problem in ramp metering research. But “multum in parvo”, and its technical complexity and implementation cost are much lower than coordinated control. Moreover, in quite a lot of cases, the isolated ramp metering strategy is more effective than the coordinated strategy, so it is of important value of research.

3.1 Fixed-Time Strategies

As early as 1965, Wattleworth proposed the method of fixed-time control of the on-ramp [3]. In the absence of modern detection technology and computer technology, this method developed a ramp metering strategy based on mathematical programming theory, with the traffic capacity as constraint, and with the maximum of the mainline traffic volume as an objective function. Since then, the prelude of the research and application of ramp metering has been opened [4].

The fixed-time strategy proposed by Wattleworth is simple for modeling, easy to implement, and also with the low implementation cost. However, it is an ideal mathematical model based on constant historical demands instead of real-time traffic volume. Later, many scholars improved it to be a multiple interval fixed-time control method [5, 6]. However, whether it is pre-timed control, time-of-day control, etc., it is formulated as a static optimization problem probably inconsistent with real traffic conditions.

3.2 Responsive Strategies

Responsive control strategy considers the dynamic conditions of traffic demand, which mainly depends on the technological level of electronic and embedded system. It can be divided into three categories: demand-capacity control method, density control method (or occupancy control strategy) and gap-acceptance control method [7].

Demand-Capacity Control Method

In 1975, Masher [8] proposed a single ramp demand-capacity control strategy with the traffic volume as the control parameter, and it adjusts the ramp traffic volume entering the mainline in order to control the mainline traffic volume not to exceed the downstream traffic capacity, thus to fully utilize the mainline.

demand-capacity control method expressed by the downstream traffic capacity

Merely Considering the difference between section capacities, the ramp metering rate (r) only depends on the upstream traffic flow (q_u) and the downstream capacity (C_d). The former can be obtained by detectors, and the latter is fixed. The basic calculation method is:

If $C_d > q_u + r$, the ramp is open and there is no control;

If C_d is slightly larger or equal to $(q_u + r)$, ramp metering rate $r = C_d - q_u$;

If $C_d < q_u + r$, the ramp is closed.

Where C_d is a constant parameter, the input parameter is q_u , and the output parameter is r . It considers only the physical capacity of the road, and is a non-feedback control with single input and single output. This strategy can be adopted in the area of low density.

demand-capacity control method combined with downstream occupancy factors

In combination with the occupancy rate (or density) factor, $O_{out}(k)$ is used to indicate the detection value of occupancy rate of the downstream road segment, and O_{cr} is the optimal occupancy rate, $C_d = q_{cap}$, $q_u = q_{in}$, and the basic demand-capacity control is improved [9]. It is an improvement of DC, but not clearly stated how to determine the minimum adjustment amount. If the downstream road volume is less than its capacity, then the remaining capacity of the downstream is positive, otherwise it is negative. If the remaining capacity of the downstream is negative, it indicates that the traffic volume of this expressway section exceeds its capacity, so the minimum adjustment rate should be adopted.

$$R(k) = \begin{cases} q_{cap} - q_{in}(k-1), & O_{out}(k) \leq O_{cr} \\ r_{min}, & \text{others} \end{cases} \quad (1)$$

D-C control method is an idealized strategy, easy to implement, and is widely used in North America.

Occupancy Control Method The D-C control is actually an open loop control, and has poor anti-interference ability, so the occupancy control method emerged, and the most classic method is ALINEA [9, 10], which keeps the downstream occupancy rate of the ramp less than a certain value (occupancy threshold), by detecting the downstream occupancy rate in real time and adjusting the ramp metering rate. The specific formula is as follows:

$$r(k) = r(k-1) + K_R[\hat{o} - O_{out}(k-1)] \quad (2)$$

- Where $r(k)$: the ramp metering rate at the k th time, the vehicle/h;
- K_R : the adjustment rate parameter, the vehicle/h;
- $O_{out}(k-1)$: the downstream road segment occupancy rate at the $k-1$ time;
- \hat{o} : the road segment setting occupation rate.

Obviously, ALINEA does not depend on any specific traffic flow model and any traffic conditions, and it also changes the metering rate in a relatively gentle way. Papageorgiou et al. [11] compared the ALINEA with other methods, and showed that ALINEA had its priority in application, and was widely used in many cities, such as Paris, Amsterdam, Glasgow, Munich, etc.

However, ALINEA does not take into account factors such as queuing length. So a number of improved models have emerged. AD-ALINEA [12] considers the occupancy rate to be affected by traffic composition, weather conditions, etc., and has a high sensitivity, and predicts the occupancy in real time. When it is difficult to detect downstream occupancy, UPALINEA [13] achieves ALINEA control method by detecting only

upstream traffic volume and converting to ALINEA control method with formula. X-ALINEA [13] considers the queuing situation of the ramp caused by the ALINEA algorithm, and adjusts the control rate when the maximum queue length is reached.

3.3 Adaptive Control Strategies

Adaptive dynamic programming is a cross-disciplinary field generated by the fusion of artificial neural network, optimal control and reinforcement learning, and can also be considered as the extension of reinforcement learning in discrete fields to the continuous fields [14].

The adaptive control method in the domain of transportation is to regard the traffic system as an uncertain system, feedback is used to realize dynamic optimization adjustment of signal timing. Refer to ramp metering, there are also many researches of adaptive dynamic programming, such as neural network [15], fuzzy logic [16], adaptive dynamic programming [17], genetic algorithm, particle swarm algorithm [18] and their fusion or improved form.

3.4 Integrated Optimization Control Strategy

Integrated optimization control considers the factors together such as intersection channelization, signal timing, network route decision-making, variable lane, turn restrictions and transit priority. So it is a multi-objective nonlinear programming problem; a representative method is the lane-based optimization method. The lane-based optimization method was first proposed by Wong CK, Wong SC [19, 20], which defines all design variables based on lanes, and makes a series of constraints into linear system of equations, thus ensures to be a solvable optimal model. Thereafter, the lane-based optimization method is further developed and expanded [21, 22]. Jing Zhao [23] gives an integrated signal optimization control model that considers the flexible division of the lane function. As there are more and more factors considered, the model gets closer to the actual situation, but it becomes more and more complicated.

4 Coordinated Ramp Metering Strategies and Algorithms

4.1 Coordinated with Surface Road Control Systems

As urban space resource is extremely limited, on-ramps (or off-ramps) are often very close to the ground intersection, so it is necessary to cooperative ramp metering with surface road control systems.

Kwon (2002) proposed to coordinate on-ramps and surface road intersections by defining the degree of congestion in order to balance the on-ramp congestion and the congestion of the ground intersection [24]. Changliang Y, Honghai L (2010) analyzed the composition and relationship of queuing congestion within the control range of off-ramp and auxiliary roads, and provided guidance and basis for formulating coordinated control strategies by giving the relationship diagrams of queuing congestion with flow and signal control [25]. Zhang X et al. (2013) proposed a speed control method for

ramps of expressway and took the on-ramp of Wuning Road in Shanghai as an example, the implementation effects of uncontrolled, timing control and speed control were also simulated and analyzed [26].

Most of other current studies focused on the improved ALINEA method of coordinating control systems of ramps and ground intersections.

4.2 Coordinated with Other Control Systems

With the lengthening of the main line of the expressways, other methods besides ramp metering become gradually popular, so there exist various control methods such as the mainstream speed limit, on-ramp speed limit, route guidance, etc., which can prevent congestions cooperative with ramp metering [27]. Mainstream Traffic Flow Control (MTFC) is the typical representative, which is formed by the combination of main line speed limit and ramp metering, and also coordinated control of ramp metering with route guidance, coordinated control of ramp metering and monetary cost, and so on. Merely ramp metering sometimes can't prevent traffic congestions effectively, and may even cause the queue length of the ramps to be too long, which may also make the congestions of the auxiliary roads and adjacent intersections. So the combination of ramp metering and route guidance measures (such as VMS (variable message signs), vehicle navigation systems, radio, etc.) to adjust the spatial distribution of traffic, and to guide them to pass through uncongested roads and intersections.

In brief, ramp metering gradually works coordinated with other control systems instead of independently.

4.3 Multi Ramp Joint Control Strategy

Multi ramp joint control strategy focuses on metering several ramps or all the ramps of the whole systems, it has a clear control objective, which is usually travel time or traffic flow of the entire system, by considering system constraints, such as maximum allowable queue length, bottleneck capacity, and so forth. However, these algorithms' performances basically depend on the quality of input data (such as OD matrices, estimated ramp capacity, and predicted traffic demands), and the traffic models used [2].

Bottleneck Algorithm

Jacobsen et al. proposed the Bottleneck algorithm in 1989 [28]. As one of the competitive ramp metering algorithms: At the single-point level, it uses the occupancy control method to obtain the regulation law of each ramp. At the cooperative level, it first detects and determines the mainstream bottleneck, and then use the principle of flow conservation to calculate the total adjustment values of upstream on-ramps, and finally after comparing the two adjustment values at each ramp, it takes the smaller value as the actual adjustment value. This strategy is simple and easy, and it is essentially identical with the hierarchy control that appears later. The differences between them lie in the complexity of the upper-level target optimization method. Many scholars have also applied Bottleneck for ramp metering research and got good results.

Linear Programming Method (LP) and Nonlinear Programming Method Yoshino et al. [29] first developed and implemented the algorithm on the Hanshin Expressway in Japan. The classical system optimal control can be divided into two types: linear open loop control and nonlinear open loop control. Linear open loop control mainly refers to LP control. AMOC [30] (advanced motorway optimal control), DMCS [31] (dynamic metering control system), and ARMS [32] (advanced real-time metering system) are nonlinear open-loop control.

The LP control method proposed by Yang [33] considered the constraints of ramp queuing, and analyzed the characteristics, its model is as follows:

The total mileage of the inflow traffic (car · km) is the largest

$$\max \sum_{i \in I} \bar{d}_i^s U_i^s \tag{3}$$

Mainline section capacity constraint

$$\sum_{r \leq s} \sum_{i \in I} Q_{ik}^{rs} U_i^r \leq C_k^o, \forall s, \forall k \tag{4}$$

Constraint of off-ramps queuing

$$\sum_{r \leq s} \sum_{i \in I} Q_{ik}^{rs} U_i^r \leq C_j^o + Z_j^o - Y_j^s, \forall s, \forall j \tag{5}$$

Constraint of on-ramp queuing and demand

$$D_i^s - (L_i^o - l_i^s) \leq U_i^s \leq D_i^s + L_i^s, \forall s, \forall i \tag{6}$$

Where: U_i^s the on – ramp i vehicles entering the mainline from during the s unit; \bar{d}_i^s the mean mileage of the inflow traffic from the on-ramp i during the s th unit; Q_{ik}^{rs} the proportion of traffic from the i th on-ramp within the r th unit control period to the j th off-ramp within the s th unit; C_k^o the capacity of the k th section of the main line; C_j^o the maximum number of allowed queued vehicles; Z_j^o the maximum queuing vehicles' numbers of the j th off-ramp; Y_j^s at the initial of the s th unit control period off-ramp j queuing vehicles' numbers; D_i^s on-ramp i arrival vehicles numbers during the s unit; L_i^o the maximum queuing vehicles' numbers of the i th on-ramp; L_i^s at the initial of the s th unit the i th on-ramp queuing vehicles' numbers.

Hierarchy Control Method

In theory, the system optimal control strategy can make the expressway traffic flow always in the optimal state, but its shortcomings are obvious: the model is often too complicated, the solving process is cumbersome, and it may produce no optimal solution or convergence. Due to the particularly serious problem of the nonlinear method, the hierarchy control method proposed by Papageorgiou [34, 35] solved this problem to some extent. The aforementioned heuristic coordinated control strategy based on the qualitative semi-quantitative design idea: the bottom layer is generally implemented by an isolated ramp metering algorithm, and the coordinated layer is designed by experiences. The hierarchy control strategy differs in that there is a

relatively obvious control objective function in the coordinated layer, but the solution does not require a complete mathematical algorithm for the derivation and demonstration, which can be regarded as a transition between heuristic coordinated control strategy and LP (and NLP). And recently there are some studies about the optimal controller method in the hierarchy control architecture [36, 37].

METALINEA

METALINEA [38] is a type of ALINEA application in coordinated control and belongs to the Linear Quadratic (LQ) control type.

$$r(k) = r(k-1) - K_1[o(k) - o(k-1)] - K_2[O(k) - \hat{O}(k)] \quad (7)$$

Where: $r = [r_1 \cdots r_m]^T$: the vector of m controllable on-ramp volumes; $o = [o_1 \cdots o_m]^T$: the vector of n measured occupancies on the expressway sections; $O = [O_1 \cdots O_m]^T$: the vector of m measured occupancies of the downstream mainline; $\hat{O} = [\hat{O}_1 \cdots \hat{O}_m]^T$: the vector of m objective occupancies of the downstream mainline; $K_1 \in R_{m \times m}$, $K_2 \in R_{n \times n}$: Gain matrix.

4.4 Network Flow Coordinated Control

The transportation network is a combination of various concerned systems, and the application of network flow control theory on ramp metering is widely studied.

Aiming at the expressway congestion problem, a network flow control method proposes ramps as control buffer and mainstream as mainline, by communication network theory, applies packet congestion control and active queue management control to establish the control system structure of the expressway network. Oh JS et al. [39], Miller HJ [40] and other researchers portrayed the interchange space structure model [41, 42]. For the first time, Li [43] proposed a new abstract network description form for the interchange. This general form is not only a plan, but also shows the basic topology of the interchange, and it is also a complex and weighted network, which indicates the spatial connection structure and physical characteristics for the interchange traffic flow transformation.

5 Conclusion and Outlook

In summary, researches on ramp metering of urban expressway are very abundant, which shows ramp metering is an effective and essential measures to solve expressway problems. This paper discusses several isolated and coordinated ramp metering strategies and algorithms, from multi-target coordinated control of isolated ramp to multi-ramp coordinated control of one-direction mainline, then to regional coordinated control of multi-ramp. As the factors considered increase, the algorithms are closer to the actual traffic conditions. In Future ramp metering researches will focus on integration and interaction with ground transportation systems and other ITS systems (including Advanced Vehicle Control Systems, Dynamic Route Guidance System and

Advanced Traveler Information System). However, it requires more inputs mainly based on real-time data, so also a challenge for SCADA system.

Acknowledgments. This research is supported by the Shanghai Science and Technology Committee (Number 17DZ1204003), and by Key Project of Special Development Fund for Zhangjiang National Independent Innovation Demonstration Zone in Shanghai (Number 201705-JD-C1085-056, 207105-JD-C1085-072).

References

1. Shaaban, K., Khan, M.A., Hamila, R.: Literature review of advancements in adaptive ramp metering. *Procedia Comput. Sci.* **83**, 203–211 (2016)
2. Wang, H.: Advanced control strategy of urban expressway traffic flow. Beijing University of Technology (2003)
3. Wattleworth, J.A., Berry, D.S.: Peak period analysis and control of a freeway system some theoretical investigations. *HRB Rec* **89**, 1–25 (1965)
4. Yang, J.: Research on ramp traffic mechanism and ramp metering design of urban expressway. Chongqing Jiaotong University (2008)
5. Iida, Y., Hasegawa, T., Asakura, Y., et al.: A formulation of onramp traffic control system with route guidance for urban expressway. In: IFAC 6th International Conference on Control in Transportation Systems. France, pp. 229–236 (1989)
6. Iida, Y., Asakura, Y., Tanaka, H.: Optimal on-ramp traffic control method for urban expressway network with multiple routes. In: *Proceeding of Infrastructure Planning*, pp. 305–312 (1986)
7. Zheng, J., Dong, D., Chen, H.: Comparison analysis on ramp metering strategies for urban expressway. *Comput. Meas. Control* **14**(2), 196–199 (2006)
8. Masher, D.P., Ross, D.W., Wong, P.J., et al.: Guidelines for design and operating of ramp metering systems. Stanford Research Institute, Report NCHRP3-22, SRI Project 3340 (1975)
9. Papageorgiou, M., Blosseville, J.M., Hadj-Salem, H.: Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris, Part II: coordinated on-ramp metering. *Transp. Res. Part A (General)* **24**(5), 362–370 (1990)
10. Papageorgiou, M., Hadj-Salem, H., Blosseville, J.M.: ALINEA: a local feedback control law for on-ramp metering. *Transp. Res. Rec. J. Transp. Res. Board* **1320**, 58–64 (1991)
11. Papageorgiou, M., Hadj-Salem, H., Middelham, F.: ALINEA: a local ramp metering, summary of field results. *Transp. Res. Rec. J. Transp. Res. Board* **1603**, 90–98 (1997)
12. Papageorgiou, M.: Traffic control. In: *Handbook of Transportation Science*, 2nd edn. Kluwer Academic Publishers, Boston (2002)
13. Smaragdis, E., Papageorgiou, M.: A series of new local ramp metering strategies. In: 82nd Annual Meeting of the Transportation Research Board, Washington DC, Paper No. 03 3171 (2003)
14. Zhao, D., Liu, D., Yi, J.: Review of urban traffic signal optimization control method based on adaptive dynamic programming. *Acta Automatica Sinica* **35**(6), 676–681 (2009)
15. Zhang, H.M., Ritchie, S.G.: Freeway ramp metering using artificial neural networks. *Transp. Res. Part C (Emerg. Technol.)* **5**(5), 273–286 (1997)
16. Chen, L.L., May, A.D., Auslander, D.M.: Freeway ramp metering using fuzzy set theory for inexact reasoning. *Transp. Res. Part A (General)* **24**(1), 15–25 (1990)

17. Xu, J., Wang, F.Y.: Ramp metering based on adaptive critic designs. In: Proceedings of IEEE International Conference on Intelligent Transportation Systems, Toronto, pp. 1531–1536 (2006)
18. Teodorovic, D.: Swarm intelligence systems for transportation engineering: principles and applications. *Transp. Res. Part C (Emerg. Technol.)* **16**, 651–667 (2008)
19. Wong, C.K., Wong, S.C.: Lane-based optimization of signal timings for isolated junctions. *Transp. Res. Part B* **37**, 63–84 (2002)
20. Wong, C.K., Wong, S.C.: Lane-based optimization method for multi-period analysis of isolated signal control junctions. *Transportmetrica* **2**(1), 53–85 (2006)
21. Zhao, J., Ma, W., Zhang, H., et al.: Increasing the capacity of signalized intersections with dynamic use of exit lanes for left-turn traffic. *Transp. Res. Rec. J. Transp. Res. Board* **2355** (2355), 49–59 (2013)
22. Sun, W., Wu, X., Wang, Y., et al.: A continuous-flow-intersection-lite design and traffic control for oversaturated bottleneck intersections. *Transp. Res. Part C Emerg. Technol.* **56**, 18–33 (2015)
23. Zhao, J., Liu, Y.: Integrated signal optimization and non-traditional lane assignment for urban freeway off-ramp congestion mitigation. *Transp. Res. Part C Emerg. Technol.* **73**, 219–238 (2016)
24. Kwon, T.M., Ambadipudi, R., Kim, S.: Signal operations research laboratory for development and testing of advanced control strategies, Phase II. *Center Transp. Stud.* 55–70 (2002)
25. Changliang, Y., Honghai, L.: Coordinated control strategy of expressway-exit-auxiliary road signal and urban road intersection signal. *Road Traffic Saf.* **10**(3), 38–44 (2010)
26. Zhang, X., Pan, C., Xun, Y.: A new signal coordination control model of joint area at the expressway conventional network. In: *LISS 2012* (2013)
27. Lan, L., Jian, S., Keping, L.: On-ramp speed control of urban expressway. *J. Transp. Inf. Saf.* **29**(3), 15–19 (2011)
28. Jacobsen, L., Henry, K., Mahyar, O.: Real-time metering algorithm for centralized control. *Transp. Res. Rec.* **1232**, 17–26 (1989)
29. Yoshino, T., Sasaki, T., Hasegawa, T.: The traffic-control system on the Hanshin expressway. *Interfaces* **25**(1), 94–108 (1995)
30. Chen, O., Hotz, A., Ben-Akiva, M., et al.: Development and evaluation of a dynamic ramp metering control model. In: *8th IFAC Symposium on Transportation System*, Chania, pp. 1162–1167 (1997)
31. Liu, J.C., Kim, J., Chen, Y., et al.: *An Advanced Real Time Metering System: The System Concept*. Department of Transportation, Texas (1993)
32. Liu, J.C., Kim, J., Lee, S., et al.: The advanced distributed ramp metering system (ARMS). In: *Workshop on Parallel & Distributed Real-Time Systems* (1994)
33. Yang, X.: *Coordinated responsive metering methodologies and research of traffic system for urban expressway*. Tongji University (1996)
34. Papageorgiou, M.: A hierarchical control system for freeway traffic. *Transp. Res. Part B (Methodological)* **17**(3), 251–261 (1983)
35. Papageorgiou, M.: Application of automatic control concepts to traffic flow modeling control. In: Balakrishnan, A.V., Thoma, M. (eds.) *Lecture Notes on Control and Information Sciences*, pp. 59–162. Springer, Heideberg (1983)
36. Zhang, G., Wang, Y.: optimizing coordinated ramp metering: a preemptive hierarchical control approach. *Comput. Aided Civ. Infrastruct. Eng.* **28**(1), 22–37 (2013)
37. Meshkat, A., Zhi, M., Vrancken, J.L.M., et al.: Coordinated ramp metering with priorities. *Intel. Transport Syst. IET* **9**(6), 639–645 (2015)

38. Taale, H., Schouten, W.J.J.P., Koote, J.V.: Design of a coordinated ramp-metering system near Amsterdam. In: International Conference on Road Traffic Monitoring & Control. IET (2002)
39. Oh, J.S., Cortes, C.E., Jayakrishnan, R., et al.: Microscopic simulation with large-network path dynamics for advance traffic management and information systems. Department of Civil Environmental Engineering and Institute of Transportation Studies, University of California Irvine, UCI-ITSTS-WP99-11, pp. 6–9 (1999)
40. Miller, H.J., Shaw, S.-L.: Geographic Information Systems for Transportation Principles and Applications, pp. 55–62, 321–324. Oxford University Press, Inc., Oxford
41. Zhang, J.: Study on control and guidance method for expressway communication. China Agriculture University (2005)
42. Chen, T.: Research on models of congestion forecast and control based on systems science theories. Southeast University (2005)
43. Li, J.: Theory and Method on Ramp Traffic Control. Beijing Jiaotong University Press, Beijing (2013)



Research on the Application of TOD Theory—Taking Huijin Road Station of Shanghai Rail Transit Line 17 as an Example

Jicheng Huang¹, Xiaoqing Zeng², Yizeng Wang², Yining Chen²,
Nixuan Ye², and Zhongzheng Ma^{1(✉)}

¹ Shanghai Metro Line 17 Development Corporation, Shanghai, China
huangjicheng@sina.com, mazhongzheng2010@163.com

² College of Transportation Engineering, Tongji University, Shanghai, China
zengxq@tongji.edu.cn, felis.wangyz@foxmail.com,
2959244421@qq.com

Abstract. As a new way of urban development, TOD mode provides a new way to solve the problems in current urban construction. Shanghai Rail Transit Line No. 17 Huijin Road Station integrated the TOD concept into the construction and made adjustments and improvements in light of the actual situation. This paper studies and analyzes its application model in order to provide reference for the extensive practice of the future TOD model.

Keywords: TOD theory · Rail transit · Huijin Road Station · Urban construction

1 Shanghai Rail Transit Line 17 TOD Development Technology Innovation

At the beginning of the construction of Rail Transit Line 17, we actively adhered to the concept of comprehensive development of docking stations, and realized “communication integration”, “space integration” and “functional integration” to promote the docking construction of the subway and the city.

In the “communication integration”, we will consider all kinds of transportation modes around the Huijin Road site on Line 17, and build new transportation hubs to achieve “seamless transfer”, improve passenger transfer experience, and bring great convenience for transfer trips.

In the “spatial integration”, Line 17 is committed to the integration of above-ground and underground space. Through the adjustment of the docking form of the entrance and exit, the station space and the commercial space are integrated with each other. Through the docking of subway functions, the surrounding development and construction of facilities and equipment in the subway will be combined to ensure the harmonious appearance of the architectural landscape along the urban roads.

In the “comprehensive function”, Line 17 introduces the concept of comprehensive urban land during the construction of the depot, comprehensively develops the land

of the depot, and builds new commercial, residential and park buildings to form a super-large-scale composite urban block, changing the past. The phenomenon of low land utilization in the depot.

1.1 Shanghai Rail Transit Line 17 Development Concept

Shanghai Rail Transit Line 17 is positioned to serve the suburban line of Qingpu District, and is also the city express line connecting the Qingpu New Town. The entire Line 17 runs along the east-west direction. It is a radiation radiating from the central city to Qingpu District. It runs through the passage of the Qingpu District East-West Passenger Corridor. The Zhujiajiao, Qingpu New City, Zhaoxiang New Town and Zhaoxiang Commercial and Business District are connected in series. Xujing Town, Huaxin supporting commercial housing base, China Expo, Hongqiao Hub and other important areas.

As the node of the rail transit line, the station is the distribution point for the citizens to take the rail transit. According to the design concept of the combination of Line 17 and surrounding development, the mode of the station should be closely coordinated with the urban public space, so that the station and the upper building are integrated, forming a three-dimensional urban space combining above and below ground, becoming the relative center of the urban area. A multi-functional composite urban complex. The stations in specific areas of the city will be considered as urban public spaces to create a real urban development gathering place.

The overall development concept of the system planning along the route is to introduce the overall development model of “Metro + Property” into the construction of Line 17. Development along the route from planning to design, from construction to management, follow the strategy of overall planning, overall development, and phased development, so as to maximize the value-added effect along the line, and use the traffic-oriented development (TOD) as a regional development model, using various policies. And the planning orientation attracts the population to move in, gathers popularity, and cultivates the passenger flow of Line 17.

2 Analysis of TOD Application in Huijin Road Station

The construction of Huijin Road Station on Line 17 fully embodies the functions of rail transit to meet the needs of urban construction and rail transit to shape the urban form. Huijin Road Station is closely integrated with the surrounding development, seamlessly connected, and basically implemented simultaneously. Through the sunken plaza, the construction of ancillary facilities, the integration of “traffic” resources, and the integration of “space” inside and outside, the three-dimensional development of above-ground and underground urban spaces is formed.

2.1 Descended

The sunken plaza completely opens the intersection of the station floor and the commercial basement and processes it into a sunken plaza. This is an ideal form of entrance and exit, which not only improves the building environment on the ground floor, but also increases the interest and ease of memorization. It is also very helpful for guiding a

large number of people in an emergency. The sinking square effectively grounds the underground space, and the indoor space is outdoorized, realizing the mutual penetration of indoor and outdoor space, which not only enriches the space level, but also improves the underground closed inward space environment (Figs. 1 and 2).



Fig. 1. Schematic diagram of the location of the sinking square at Huijin Road Station



Fig. 2. Huijin Road Station sinking square renderings

2.2 “Traffic” Resource Synthesis

The traffic function is the most basic function of the comprehensive development of rail transit stations. All kinds of transportation modes near the subway station are considered in a unified manner, which is convenient for transfer and realizes “seamless transfer”.

For example, close to the subway station to set up bus hubs, intermediate stations, “P+R” public parking lots, etc., to expand the radiation range of subway stations. The comprehensive development design of the subway should fully consider the above requirements, optimize the layout of the entrance and exit, optimize the transfer path, and realize a spacious and convenient, all-weather “seamless connection” path to enhance the passenger transfer experience.

Located at the east end of Qingpu New Town, Huijin Road Station, the surrounding area of the site is planned to be the new city public service center. From the station position, it is the “P+R” portal function of the Qingpu New City docking main city. Therefore, in addition to the planning functions of commercial, office and residential, the planning and construction of the site around Huijin Road Station has built a regional public transportation hub, including 13 bus lines, between the Yuya Road and Yuquan Road. The block is equipped with 1000 parking spaces for social vehicles, which can be connected to the “P+R” function in later operations (Fig. 3).

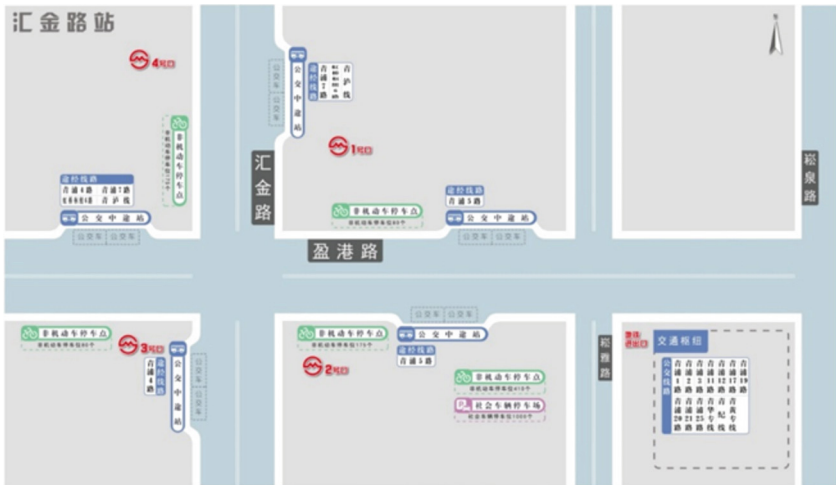


Fig. 3. Huijin Road Station public transportation facilities layout map

2.3 Integration of Internal and External “space”

The construction of the subway station is a golden opportunity for the development of underground space. Through the link effect of the rail transit station, the integration of the above-ground new underground and underground space is realized in the architectural space form. The station space and the commercial space penetrate and merge with each other to reach various spaces. “No sense of transition.”

Subway and development of comprehensive and integrated development and design, combined with the functional requirements of the subway station, docking content focuses on the following two points:

2.3.1 Docking of Entrances and Exits

When the passenger flow of the station is at the entrance and exit station, especially when leaving the station hall, the passengers can choose the route according to their own destinations: they can leave directly, that is, they can directly reach the outdoor space; they can enter the transfer passage and transfer with other modes of transportation. The above two points are the fast and effective appeal of the subway station as the most important traffic flow guidance for the traffic building; it can also enter the commercial or hotel, office and other functional spaces through the interface between the subway and the development. In the design, while ensuring the orderly organization of the three types of passenger flow, we should also consider some issues such as fire protection, civil air defense, and operation management. The subway space carriers involved in this part are mainly non-paid areas for entrances and exits and station halls.

Set up a sunken plaza, the subway entrance and exit can be directly solved in the square, which can break the sense of closure of the underground space and subtly connect the underground, ground space and entrances and exits.

The transition through the outdoor open space of the sunken plaza is a more advantageous treatment. This way, the interface and space transition between the subway and the development are natural and coherent, which can well meet the independent division of fire and civil air defense. A clear and clear operational management interface is achieved. On the other hand, it is also possible to create more underground spaces such as subway stations and commercial and transportation interchanges. Through the introduction of the sunken plaza, natural light and landscape will follow, creating a comfortable, humanized and well-oriented walking environment. The transition between spaces is more organic, integrated and safe.

In addition, in the appearance of the subway entrance and exit of the development docking, through the coordinated design of the overall style of the development building, the architectural form is more coordinated and unified, avoiding the incompatibility between the unified entrance and exit of the whole line and the various development buildings (Figs. 4 and 5).



Fig. 4. Metro Huijin Road Station Exit 1 docking Baolong Square



Fig. 5. Metro Huijin Road Station Exit 2 docking Baolong Square

2.3.2 Docking of Subway Functions

The number of subway wind shafts is large and the volume is large. At the same time, there are a large number of ground equipment, such as cooling towers and VRV outdoor units. These facilities and equipment need to occupy a considerable amount of land for outdoor installation, and the shape is ugly. It has always been designed by subway stations. The landscape “disaster area” has a considerable impact on the urban landscape.

The integration of the above-mentioned equipment and facilities into the development building through design docking is an effective countermeasure to solve the above problems, which can solve the negative impact of the “large size, large number, scattered layout” of subway facilities on the urban landscape, thus ensuring the urban roads. Coordination and unification of architectural landscapes along the route (Fig. 6).

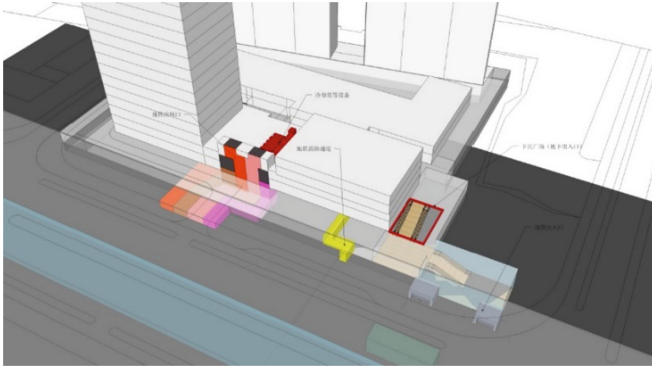


Fig. 6. The No. 2 wind shaft and cooling tower of Huijin Road Station are connected with the development building.

References

1. Wang, G., Zeng, X., Yuan, T. (eds.) Study on the influence of train control system on service quality of rail transit. In: International Conference on Service Systems and Service Management (2017)
2. Jian, L., Xiao-Qing, Z., Tuo, S., et al.: Design of the fault recording function in a railway signal microcomputer monitoring system. In: Proceedings of the CICTP 2016 Green and Multimodal Transportation and Logistics 16th COTA International Conference of Transportation Professionals, 6–9 July 2016, Reston (2016). ASCE - American Society of Civil Engineers
3. Wang, W., Zeng, X., Shen, T.: CICTP 2015: efficient, safe and green multimodal transportation. In: Proceedings of the 15th COTA International Conference of Transportation Professionals, pp. 1732–1743 (2015)
4. Zeng, X., Tao, C., Niu, Z., Zhang, K.: The study of railway control system model. In: 5th IEEE Conference on Industrial Electronics and Applications (ICIEA), June 2010



Data-Driven Safety Model on Urban Rail Transit Signal System

Yujia Chen^(✉), Xiaoqing Zeng, and Tengfei Yuan^{ID}

The Key Laboratory of Road and Traffic Engineering, Ministry of Education,
School of Transportation Engineering, Tongji University, 4800 Cao'an Road,
Shanghai, China

Chenyujia2@126.com,

{Zengxq, 15yuantengfei}@tongji.edu.cn

Abstract. The signal system is the control system in the rail transit system and is a key component to realize the safety of train operation. As more signal systems use complex computer systems instead of traditional control methods, data plays an increasingly important role in the system's work. Establishing an analytical model with data as a system unit is important for further strengthening system safety. This paper takes urban rail transit signal system as the research object, and builds a system safety model based on the data driving characteristics of the system, which provides theoretical support for further improving the safety of the new generation rail transit system.

Keywords: Data-driven · Signal system · Safety model · Urban rail transit

1 Introduction

Urban rail transit has become one of the important choices in urban travel with its characteristics of high efficiency, high energy capacity and high reliability. Safety and efficiency are the most important research topics in urban rail transit. In rail transit, there is a low coefficient of friction between the train and the track, and the rail has a fixed path, so the train has a poor ability to evade obstacles [1]. The signal system can make the train predict the obstacles in advance and avoid the collision of trains [2], which is the key link of urban rail transit safety.

In the operation of the signal system, due to the complexity and concurrency of the system, the system data faces a complex state [3]. Due to the dynamic nature of the system, the possibility of data errors is large, and the causes of the system are variable. Traditional system safety analysis focuses on the hardware and software components of the system, and the data portion of the system is rarely analyzed separately, which is common in traditional electronic component analysis. The wide application of computers makes the rail transit signal system evolve from relay logic control to computer control, which makes the role of data in the system more prominent. Correspondingly, erroneous data can also lead to dangerous behavior of the system. In order to make data work safely and effectively in the system, it is not enough to study hardware and software. It is necessary to study the data as a separate component to analyze how the system works to avoid risk [4, 5].

This paper is organized as follows. Firstly, this paper analyzes the data characteristics in the signal system and discusses its working mechanism in the system. Then, builds the data-driven system security model according to the data characteristics. Finally, the area in the signal system. For example, the controller builds the data-driven security model of the system and performs model transformation demonstration.

2 Data Characteristics

2.1 Data Component

From a computer perspective, in the physical form of existence, data is a representation of the potential state in an address in memory; in the form of existence of a software system, data represents a variable in the code. From the perspective of system function and security, the meaning of data at the software level is more practical [6], which is why more data operations are proposed in many software standards.

Traditionally, data is often distinguished by its meaning, usually with “item” as the basic unit of data. The life cycle of each “item” data includes the four processes of generation, transmission, use and extinction [7] (as shown in Fig. 1, where there may be multiple transmission and use links), indicating that a “item” of data is collected, It is passed to the final use location, then involved in calculations or control, etc., and finally functionally obsolete. For example, the speed data of a certain train at a certain time is measured by the speedometer, and then transmitted to the VOBC of the train and the ZC of the track, which are respectively used for the functions of train control and MA calculation. When the new data is obtained at the next moment, the old data is invalid. The process of. This perspective of understanding data is more common, and it is more intuitive and clear through the logical flow of data. However, data has a strong abstraction and variability that can fail at every step of the lifecycle. The complex failure of the data itself has many different forms. When the data of the whole life cycle is taken as the research unit, the complex existence form is difficult to establish contact with the system behavior [8]. Therefore, performing basic unit analysis of data in this form can make the system more complicated and difficult to analyze.

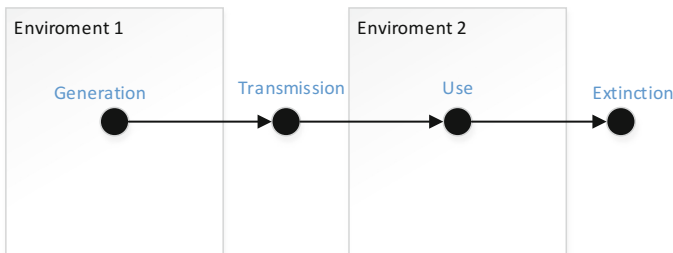


Fig. 1. Schematic diagram of the existence of data.

In this paper, based on the concept of the figure, the structure of the data “components” is used to model the structure and function of the system. Its form is similar to placeholder [9] in Tensor Flow, which treats data in the same location, the same function, the same meaning as a whole, and the “place” of the data as a component. The data value of the “place” and the like constitute its attributes. The data is cut into units in this section. On the basis of the meaning, the existing form of the data is further subdivided, which can make the system structure clearer and avoid the explosion of the state space, which is helpful for the subsequent formal analysis.

As shown in Fig. 2, the attributes of the data component are divided into identifiers, data source, values, invalid content, and so on. The component does not pay attention to the logical meaning of the data in other forms, only focusing on upstream and downstream data transfer. The topological relationship determines where data exists in the workflow. As a result of this processing, the data is no longer in the same environment as in Fig. 1, spanning several environments (the environment can refer to modules or subsystems), so that failure analysis can be performed on a single module. Taking the data component as the node, the system’s processing mode of data processing is used as the connection mode, and the data is connected to form the data component workflow model of the system to form a complete system.

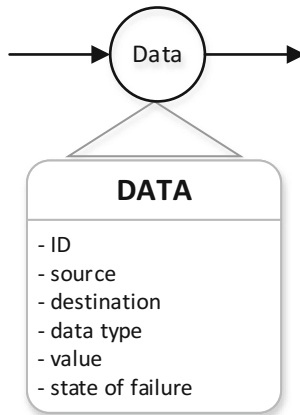


Fig. 2. Schematic diagram of the data component.

2.2 Data Safety Features

The Function of the signal system is essentially the processing of data. During the course of its work, a large amount of dynamic data undergoes processes such as generation, transmission, use and extinction [10]. Taking the regional controller (ZC) as an example, the ZC periodically receives information such as the train position, traveling direction, and train operation transmitted by the VOBC through the DCS, and

periodically transmits the calculated MA data to the VOBC; the ZC and the DSU perform information. Interaction, first complete the comparison of the database version number, and then update the database version number and the corresponding data content; the data sent by the CI to the ZC mainly includes the information of the route for the train and the state of the obstacle within the access range; ATS The information sent is mainly information such as confirmation of information sent by ZC and temporary speed limit.

Constraint Relationship. Constraint relationship refers to the mathematical connection between data objects, that is, the relationship between this data and other data can be obtained through mathematical calculation. This relationship includes both the logical relationship between the attributes of the data and the computational relationship between the data values. Due to the correctness and security requirements of the system, this constraint relationship between data is more concentrated in the security constraints.

Consistency Requirements. The input data of the system, the description of the real physical scene it characterizes, or the purpose of the data source to transmit the data must be consistent [11]. When data is generated, transmitted, and used outside the system, certain errors may occur, which may affect the normal operation of the user subsequent to the data. Normally, the correctness of the system's input data is unknown to the system. The general system can judge the legality of the data, and can also establish certain logical constraints, and strengthen the security measures according to the security level of the consistency requirements. Such as: train length data.

Data Timeliness. Due to the dynamic nature of the system, data validity will fail over time. The time validity of the input and the timeliness of the output data must be considered during the system design process. The usage time of all inputs used to generate the output instructions must be limited. Once the time limit is exceeded, it is marked as invalid. The time-dependent reflection of data is the duration of the data life cycle and the length of time that the data is effectively used by the system. The normal operation of the data life cycle is also a necessary condition for the normal operation of the system.

2.3 Data Failure Mode

As described in Sect. 2.1, the operation of data includes four processes of generation [13], transmission, use, and extinction [12]. In each process of the life cycle, data has the possibility of failure, and the data failure of each life cycle stage has its own characteristics.

Relative to the failure characteristics of the physical equipment unit and the software component unit, the failure form of the data has two characteristics: value characteristic and numerical relationship.

Value Characteristics. Data has the property of “value”, and its value varies according to the type of data. However, the influence of its value on system security may be biased. If the value is wrong, the distance between the MA and the MA is more dangerous than the distance. The high speed of the train does not affect the vehicle in front, but may affect the safety of the vehicle.

Numerical Relationship. There are numerical characteristics such as comparison relationship and calculation relationship between data, so there is a possibility of mutual influence. For example, when the vehicle speed is very low, the error of the mobile authorization may cause the emergency braking of the train, and if the vehicle speed is high, it may cause serious consequences such as collision.

Due to the numerical characteristics of the data, the failure mode of the data varies depending on the type of the value. According to the data type, define the data failure type, as shown in Table 1, so that it can meet the security evaluation requirements of the regional control system.

Table 1. Definition of failure type.

Failure type	Data type	Failure mode	Symbol	Description
No failure		Normal	*	There is no failure occurred
Numerical type	Numerical	Greater	gt	Greater than normal
		Less	lt	Less than normal
		Value	v	Value error (uncertainty vs. normal value)
	Boolean	True to False	tf	Output true to false
		False to True	ft	Output false to true
	String	Wrong String	wstr	Wrong string
	Enumerate	Others	ot	Wrong for other enumerations
	List	Order	ord	Sequential error
		Missing	miss	Intra-sequence data missing
		Insertion	ins	Insert redundant data
ItemWrong		itmw	Entry error	
Function type		omission	om	Lost output

3 System Model

3.1 Hierarchy

Most of the functions of the signal system have reached a general consensus in the industry and formed corresponding industry specifications (such as IEEE 1474, etc.). However, due to different vendors and system updates, there are many differences in

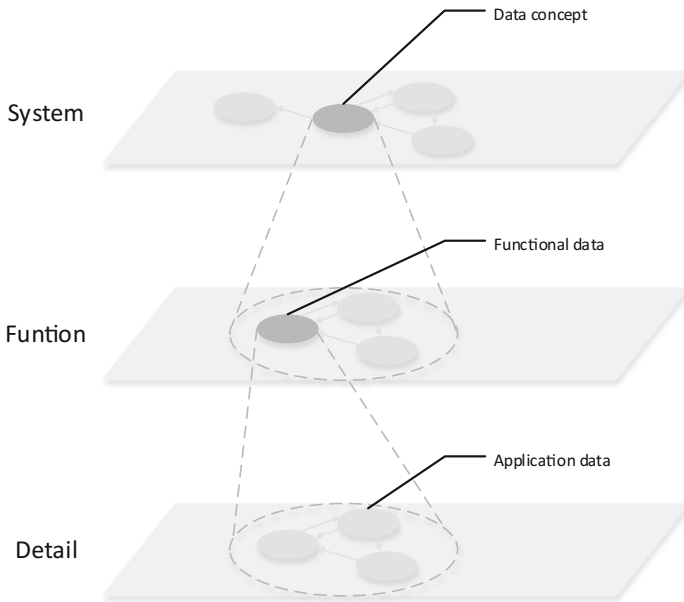


Fig. 3. System hierarchy and data.

the details of the signal system, so the data details of different systems are also very different. In the design work of the system, it generally includes the process of system requirements, function design, and detailed design. In the process of system design and development, data synchronization within the system is refined. Therefore, when performing system security analysis, it is also necessary to analyze the system's security weaknesses step by step according to the designer's description, and then propose a solution to strengthen the system safety.

The design process of the system generally includes the steps of requirements design, functional design and detailed design. As shown in Fig. 3, in the definition phase of system function and security requirements, the data structure is generally not involved, but there will be description of the data content and function. The data here is the concept data. After defining the specific functions, the function requirements are started. At the same time, the input and output contents of the function are defined, which is called function data. In the detailed design of the function, the interface is embodied and realized in consideration of the requirements, and the data is the application data that needs to be specifically used in the system.

3.2 Hierarchical Coupling

In the multi-level model, coupling problems are involved between different levels. The role of hierarchical coupling is to integrate the various components of the system to provide an analytical role. Data components are connected by functional components that need to be integrated together by a coupling algorithm to form a complete system model that can be used for analytical calculations. A component coupling algorithm is built to couple the components of the same level, and then the association of components of different layers is established to finally achieve system coupling.

For components in the same layer, the data component is connected to the functional component alternately. For different scenarios, it is necessary to first determine the order of the functional components, and then determine the specific structure of the data flow to form the target model of the analysis. The same layer component coupling algorithm is shown in Table 2.

Table 2. The same layer component coupling algorithm.

<p>• Algorithm</p> <p>1. Initialization: input data component set DcI, output data component set DcO, search direction Dir (1 indicates forward, 0 indicates reverse)</p> <p>2. $Sd \leftarrow$ (if $Dir = 1$ then DcI else DcO), Sd is data component set</p> <p>3. $Endd \leftarrow$ (if $Dir = 1$ then DcO else DcI), $Endd$ is the search-end data component set</p> <p>4. Repeat</p> <p style="padding-left: 40px;">(a) $Sf \leftarrow \bigcup_{i=1}^{length(Sd)} \left\{ \begin{array}{l} Sd_i.output_nodes, Dir = 1 \\ Sd_i.input_nodes, Dir = 0 \end{array} \right\}$, Sf is the function component set</p> <p style="padding-left: 40px;">(b) Combine (Sd, Sf)</p> <p style="padding-left: 40px;">(c) $Sd \leftarrow \bigcup_{i=1}^{length(Sf)} \left\{ \begin{array}{l} Sf_i.output_nodes, Dir = 1 \\ Sf_i.input_nodes, Dir = 0 \end{array} \right\}$</p> <p>• Until $Sd \subseteq Endd$</p>

For different layers of the model, a depth-first search strategy is used to search along the established direction according to the calculation requirements, until each search direction reaches the bottom search end. The inter-layer coupling algorithm is shown in Table 3.

Table 3. The inter-layer coupling algorithm.

<p>1 Algorithm 3.1</p>	
<p>1. Initialization: function component Df, its input data component set DcI, its output data component set DcO, Search Direction Dir(1 indicates, 0 indicates reverse)</p>	
<p>2. $Sd \leftarrow$ (if $Dir = 1$ then DcI else DcO), Sd is data component set</p>	
<p>3. $Endd \leftarrow$ (if $Dir = 1$ then DcO else DcI), $Endd$ is the search-end data component set</p>	
<p>4. Repeat</p>	
<p>(a) $Sd \leftarrow$ if $Df.isBaseNode()$ then</p>	$\left\{ \begin{array}{l} Df.output_nodes, Dir = 1 \\ Df.input_nodes, Dir = 0 \end{array} \right\}$
<p>(b) Combine (Sd, Df)</p>	
<p>(c) if not $Df.isBaseNode()$ then $Sf \leftarrow$</p>	$\left\{ \begin{array}{l} Sd.output_nodes, Dir = 1 \\ Sd.input_nodes, Dir = 0 \end{array} \right\}$
<p>(d) for each cf in Sf as Df, repeat procedure 4</p>	
<p>• Until $Sd \subseteq Endd$</p>	

4 Case Study

Take the core function of the regional controller (ZC), the mobile authorization computing function, as an example to build a data-driven system model.

During the work of the mobile authorization calculation function, ZC needs to interact with the ATS, CI, VOBC, and DSU subsystems in real time during each work

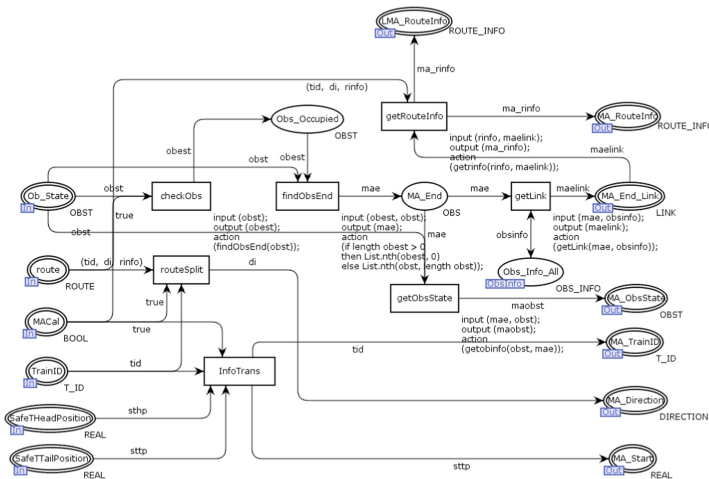


Fig. 4. Data-driven model of MA calculation Module in ZC.

cycle to obtain the latest data reflecting the train status and line status. To achieve the change and maintenance of mobile authorization. A data-driven module model is constructed based on a colored Petri net, as shown in Fig. 4.

Correspondingly, in the signal system, the corresponding important part, VOBC, responds according to the mobile authorization given by ZC, and completes the corresponding function. The data driven model of VOBC is shown in Fig. 5.

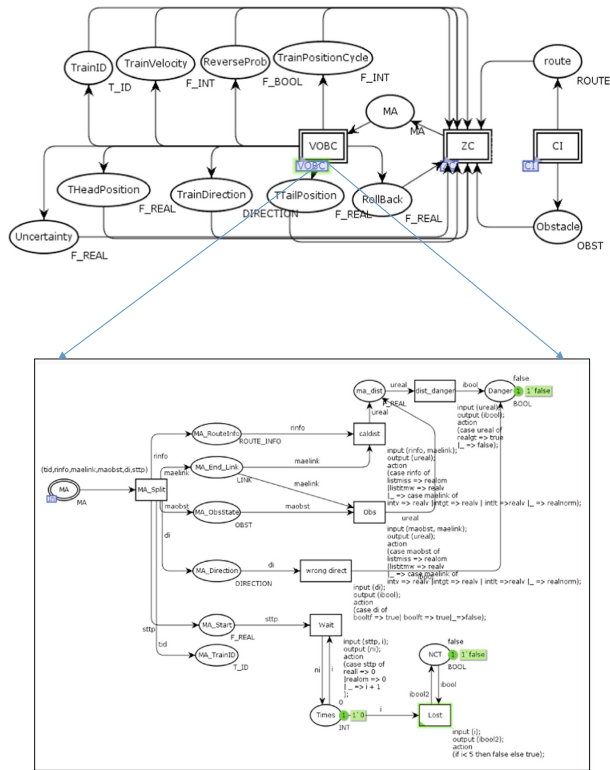


Fig. 5. Data-driven model of VOBC in MA calculation.

5 Conclusion

This paper first studies and analyzes the data characteristics of urban rail transit signal system, and then designs and constructs the security analysis model of data driven system according to its characteristics. Finally, taking the mobile authorization function in the signal system as an example, the corresponding construction is carried out by using colored Petri net. The data driven model of the subsystem. For data-driven systems, the establishment of a security analysis model for data objects is of great significance for the further strengthening of data-based urban rail transit security.

Acknowledgments. The project is supported by Tongji University and the project “Research on the Practice and Improvement of the Construction Innovation Technology of Shanghai Rail Transit Line 17” (Number JS-KY18R022-7).

References

1. Farooq, J., Soler, J.: Radio communication for Communications-Based Train Control (CBTC): a tutorial and survey. *IEEE Commun. Surv. Tutor.* **19**(3), 1377–1402 (2017)
2. Gao, C.: Communication-based train control system. *Modern Urban Transit* (2007)
3. Short, R.: Safety assurance of configuration data for railway signal interlockings. In: *Proceedings of the International Conference on System Safety*. IET (2006)
4. Harrison, A., Pierce, R.: Data management safety requirements derivation. Railtrack: West Coast Route Modernisation Internal report (2000)
5. Littlewood, B., Popov, P., Strigini, L., et al.: Modelling the effects of combining diverse software fault detection techniques. In: *Formal Methods and Testing*, pp. 345–366. Springer (2008)
6. Leveson, N.G., Harvey, P.R.: Analyzing software safety. *IEEE Trans. Softw. Eng.* **5**, 569–579 (1983)
7. Faulkner, A., Nicholson, M.: An assessment framework for data-centric systems. In: Dale, C., Anderson, T. (eds.) *Proceedings of the Twenty-Second Safety-Critical Systems Symposium*, Brighton. Safety Critical Systems Club (2014)
8. Faulkner, A., Storey, N.: The role of data in safety-related railway control systems. In: *Proceedings of the 19th Systems Safety Conference*, Huntsville (2001)
9. Girija, S.S.: Tensorflow: large-scale machine learning on heterogeneous distributed systems (2016)
10. Military Standard: Procedure for performing a failure mode effect and criticality analysis. United States military procedure MIL-P-1629 (1949)
11. Wang, H., Liu, S.: Modeling communications-based train control system: a case study. In: *Proceedings of the 2nd International Conference on Industrial Mechatronics and Automation (ICIMA)*. IEEE (2010)
12. 61508-3-2010 I: Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 3: Software requirements. Basic Safety Publication
13. Recht, J.: Systems safety analysis: failure mode and effect. *National Safety News* (1966)
14. Fenelon, P., McDermid, J.A.: Integrated techniques for software safety analysis. In: *Proceedings of the IEE Colloquium on Hazard Analysis*. IET (1992)
15. Patra, S.: Software safety assurance process for railway platform software. In: *Proceedings of the 2nd International Conference on System Safety*. IET (2007)



Simulation Data Generating Algorithm for Railway Turnout Fault Diagnosis in Big Data Maintenance Management System

Ke Cui, Maojie Tang^(✉), and Dongxiu Ou

The Key Laboratory of Road and Traffic Engineering, Ministry of Education, China School of Transportation Engineering, Tongji University, Shanghai, China
1632486@tongji.edu.cn

Abstract. Currently, the identification of turnout failures mainly depends on the use of intelligent models. However, speed-up turnouts cannot provide enough fault samples for model training. A fault diagnosis model with insufficient fault samples can cause serious safety problems due to underfitting. In this paper, we are aiming at proposing a speed-up turnout fault generation algorithm (SFGA) to address the fault-insufficient problem. The algorithm analyzes data collected from the big data management system (BDMS), and then generates 11 common faults for speed-up turnout based on Bayes Regression, Harmonic Superposition, and Fixed Constraint models. Furthermore, this paper employs derivative dynamic time warping (DDTW) to calculate similarities between simulation faults and real faults for evaluation. An experiment based on real data collected from the Guangzhou Railway Bureau in China demonstrates that all simulation faults generated by SFGA are efficient, and can be used as a training set for fault diagnosis methods of speed-up turnouts.

Keywords: Railway · Speed-up turnout · Big-Data · Simulation data generation algorithm · Fault diagnosis

1 Introduction

Although the development of high-speed railway in China has brought substantial social benefits, it also puts forward higher safety requirements for essential signal equipment such as turnout and track circuit. Aiming at ensuring the safety and reliability of the signal equipment, the railway departments keep improving the application of intelligent diagnosis system based on machine learning algorithms [1], thus putting forward two missions: exact fault diagnosis and timely fault prognosis. These two demands are based on the big data management system (BDMS) which not only receive raw data from the microcomputer monitoring system (MMS) but combine and store data. However, only a small part data stored in BDMS are fault data, and the limited and incomplete faults (fault data and corresponding types) cannot satisfy the requirement of using as a training set. Any intelligent model trained with a small number of fault samples can suffer the “undertaking” which means it diagnose or

prognose faults ineffectively and inaccurately and can cause severe safety problems. Therefore, there is an urgent requirement to afford these intelligent methods by supplying enormous fault samples.

Nowadays, two kinds of data are accessible for diagnosing, including current curves and power curves. These curves can be collected or loaded by three approaches: timely collecting from MMS, loading from BDMS and simulating by generation algorithms. Among them, the first method not only cost much time but consume excessive human resources due to the low frequency of turnout operation. For the second method, it is often difficult to find fault samples since BDMS has stored immense curves. Even if the fault data are accessible in BDMS, these faults usually belong to different turnouts and are challenging to be unified into the same fault training set. Therefore, it is meaningful to obtain enough fault samples by generating simulation fault data based on particular substantial samples, including normal data and faults, and corresponding features. Actually, a large number of industries, such as metallurgy and conveyor systems have equipped with appropriate algorithms for simulation fault generation, but these methods are all related to principles or models of that fields. Therefore, this paper proposes the speed-up turnout fault generation algorithm (SFGA), which can generate significant and comprehensive faults based on the massive historical data collected from BDMS. SFGA can learn the characters for different speed-up turnouts, thus proving a practical approach for solving the problem of insufficient fault samples and evaluating relevant diagnosis algorithms.

2 Speed-up Turnout

2.1 Action Processes of Speed-up Turnout

At present, these current and power curves stored in BDMS or MMS both are sampled with a sampling period of 40 ms. Although both two kinds of curves can reflect the running state of the turnout, current curves are more stable and widely used than the other because of current curves have more evident characteristics than power curves. Thus, scholars prefer employing current curves for fault diagnosis and prognosis to power curves, and no exception is this paper [2].

According to the data characteristics of the current curve of the speed-up turnout, the switching process of the speed-up turnout can be divided into the following three stages: the start stage, action stage and release stage. Figure 1 shows a fault-free curve of a speed-up turnout. The start stage ($0 - T_1$) exhibits a peak current when the machine begins to operate; the action stage ($T_1 - T_2$) is relatively smooth, and it corresponds to the switch process of the turnout; and the release stage ($T_2 - T_3$), which is typically called the “small step”, indicates that the turnout has finished switching and has connected the relevant circuit.

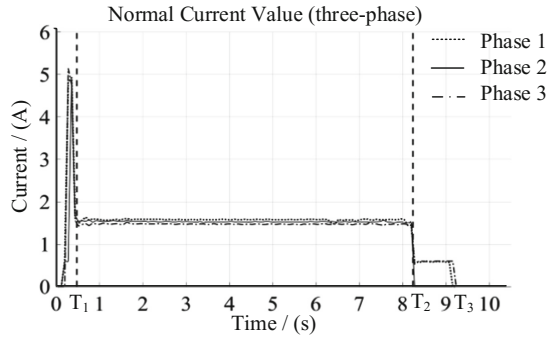


Fig. 1. Segmented normal current curve

2.2 Fault Data Classification of Speed-up Turnout

At present, there is no unified standard for the classification of fault data for speed-up turnout. In this paper, according to the existing research [3] and experience collected from railway staff, fault data of speed-up turnout are divided into 11 types: overlong release time of the relay 1DQJ, machine idling, abnormal fluctuation, poor contact in the action circuit, conversion current exceeded, electric relay 2DQJ turnout failure, release failure, turnout blocked, lock failure, blocking in the gap and open start-up circuit. These faults referred to as $M_1 - M_{11}$, are described in Table 1.

Table 1. Fault types and corresponding attributes

Symbol	Fault types	Corresponding curve characteristics	Fault mechanism
M_1	Overlong release time of the relay 1DQJ	Overly long "small steps"	Poor time characteristics of the open-phase protector
M_2	Machine idling	Overly long conversion time	Start contact cannot be disconnected
M_3	Abnormal fluctuation	Abnormal fluctuation in the action current	Oil deficiency of slide plate, or switch rail warped
M_4	Poor contact in the action circuit	Abrupt change in the action current	Indicate contact rust
M_5	Conversion current exceeded	Conversion current that exceeds the limit	Block in the conversion process
M_6	Electric relay 2DQJ turnout failure	A "small steps" curve	Open electric relay 2DQJ circuit
M_7	Release failure	Rising current in the release stage	Poor outdoor diode
M_8	Turnout blocked	Current continues to rise/fall in the action stage, no "small steps"	Foreign object exists, turnout cannot be turned into place
M_9	Lock failure	A long-term reverse curve	Some large matter exists, turnout cannot be switched in place
M_{10}	Blocking in the gap	Missing "small steps"	Indicate circuit failure
M_{11}	Open start-up circuit	Zero value curve	Open start-up circuit

In Table 1, the faults characteristics represent the curve characteristics of the fault reflected on the current curve, and the fault mechanism indicates the main reason for this kind of fault.

3 Speed-up Turnout Fault Generation Algorithm

SFGA consists of extracting data features from current curves, choosing the template curve and generating simulation faults based on corresponding templates.

3.1 Extracting Data Features from Current Curves

The durations of speed-up turnouts, in a fixed sampling period, vary from 5 s to 15 s, thus causing different dimensions for each current data. In order to unify the dimension of data, it is primary to extract features from all current curves. There are a large number of extraction methods [4–6] can cover this task, and this paper utilizes one way based on these previous studies to handle the problem at hand. The selected features are shown in Table 2.

Table 2. Features of different stages

Stage	Object	Features
The start stage	Current	Time span; Maximum; Mean; Median
The action stage	Current	Time span; Maximum; Mean; Median; Minimum; Standard deviation; Peak factor; Fluctuation factor
The slow release stage	Current	Time span; Maximum; Mean; Median; Minimum; Standard deviation; Peak factor; Fluctuation factor

In Table 2, features are selected based on curve characteristics shown in Table 1. In the start stage, only four features are chosen since nearly no fault can occur in this stage; In action stage and the release stage, basic statistic features, such as medium and mode, and fluctuation features for evaluating smoothness of curves are selected due to the high complexity of these stages. When finishing the extraction process, the feature vector F can be generated:

$$F = [f_1, f_2, \dots, f_{20}] \quad (1)$$

where $f_1 \sim f_4, f_5 \sim f_{12}$ and $f_{13} \sim f_{20}$ represent features extracted from the start stage, the action stage and the release stage respectively.

3.2 Choosing the Template Curve from Feature Vectors

A template curve is the representative of all normal current curves, which includes four essential characteristics. Firstly, template is a kind of normal current curve and mustn't be a fault; Secondly, templates vary from different turnouts; Thirdly, templates change

with time and turnouts aging; Fourth, the distance between the corresponding feature vector F^* of the template and the feature vectors of faults should be as large as possible in the vector space Θ composed by all feature vectors.

According to the four characteristics, the template should be the cluster center of all normal data in the Θ . However, the limited current curves cannot span the Θ but the Θ^V , which is the subspace of Θ . In this paper, we employ a clustering method named K-medoids [7] to tackle the problem. The details of K-medoids are list as following.

Step 1: Generating a matrix $F_{m \times 20}$ with m rows and 20 columns by extracting features from m current curves, then select one row as the initial center F^0

$$F^0 = [f_1^0, f_2^0, \dots, f_{20}^0] \quad (2)$$

Step 2: Defining the cluster validity index: the sum of absolute differences (SAD) by Eq. (3)

$$SAD = \sum_i^m \sqrt{\sum_{j=1}^{20} (f_j^0 - f_j^i)^2} \quad (3)$$

Step 3: Selecting a row F^k from $F_{m \times 20}$ except F^0 and calculate the SAD of F^k

Step 4: Setting F^k as a new center if the SAD of F^k is less than that of F^0

Step 5: If the current SAD is the minimum after going through all rows, its corresponding row is the template $f^*(x; t)$, otherwise back to Step 3,

$$f^*(x; t) = [x_1^*, x_2^*, \dots, x_{n_1}^*, \dots, x_{n_2}^*, \dots, x_n^*] \quad (4)$$

where n_1 represents the index between the start stage and the action stage, and n_2 is the index between the release stage and the action stage.

3.3 Simulation Fault Generation

Eleven kinds of fault types are divided into three parts, i.e. time anomaly fault (M_1, M_2), curve fluctuation fault (M_3, M_4) and fixed type fault ($M_5 \sim M_{11}$).

Time Anomaly Fault. This kind of fault represents that faults are abnormal in durations, especially excessive durations. This paper aims at simulating different durations using Bayes Regression [8], which extend data in the time domain by prior knowledge. The details of Bayes Regression are introduced as follows.

For linear regression, it can be illustrated by Eq. (5),

$$y = \omega_0 + \sum_{m=1}^M \omega_m^T \phi_m(x) + \delta \quad (5)$$

where $\phi_m(x)$ is the kernel function and ω_m is the corresponding coefficient of $\phi_m(x)$. δ is noise term and is subject to normal function ($\delta \sim N(0, \beta^{-1})$).

In the ordinary linear regression, ω is unknown but a constant value. For a random sampling data $\mathbf{D}^M = \{x_i, y_i\}_{i=1}^M$, the ordinary linear regression can be equivalent to solving an optimization problem:

$$\arg_{\omega} \min \sum_{m=0}^M y_m, \omega_m^T \phi_m(x)_n \quad (6)$$

In the Bayes Regression, ω is a vector of random variables, thus shown in Eq. (7), solving ω equals to solving the maximum posterior probability with given dataset \mathbf{D} .

$$\arg_{\omega} \max P(\omega | \mathbf{D}^M) \quad (7)$$

According to the Bayesian formula,

$$P(\omega | \mathbf{D}^M) = \frac{P(\mathbf{D}^M | \omega) P(\omega)}{P(\mathbf{D}^M)} \propto \mathcal{L}(\omega | \mathbf{D}^M) P(\omega) \quad (8)$$

where $\mathcal{L}(\omega | \mathbf{D}^M)$ is the likelihood with a given dataset. $P(\omega)$ represents the prior probability of ω . Also, $P(\mathbf{D}^M)$ is the normalized parameter.

This algorithm calculates posterior probabilities with incremental learning model with an enormous data $\mathbf{D}^M = \{x_i, y_i\}_{i=1}^M$

$$P(\omega | \mathbf{D}^M) = \frac{p(x_m | \omega) p(\omega | \mathbf{D}^{M-1})}{\int p(x_m | \omega) p(\omega | \mathbf{D}^{M-1}) d\omega}. \quad (9)$$

For the incremental learning model, $p(\omega | \mathbf{D}^0) = p(\omega)$ can be set as an initial estimate at first. Then the samples are input in turn, and the posterior probability of this output is taken as the prior probability of the next input. Besides, a Gaussian distribution also is chosen to generate the $p(\omega)$ for decision makers without any experience,

$$P(\omega) = N(\omega | m_0, S_0) \quad (10)$$

where m_0 and S_0 are the mean value and the covariance of the prior distribution respectively, also, the posterior distribution satisfies normal distribution based on the principle of conjugate distribution,

$$P(\omega | \mathbf{D}^M) = \frac{\prod_{m=1}^M N(y_m | \omega^T \phi(x_m), \beta^{-1}) \times N(\omega | m_0, S_0)}{P(\mathbf{D}^M)} = N(\omega | m_M, S_M) \quad (11)$$

where

$$S_M^{-1} = S_0^{-1} + \beta \phi^T \phi \quad (12)$$

$$m_M^{-1} = S_M (S_0^{-1} m_0 + \beta \phi^T Y) \quad (13)$$

$$\phi \equiv \begin{bmatrix} \phi_0(x_1) & \phi_1(x_2) & \cdots & \phi_{M-1}(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \cdots & \phi_{M-1}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(x_N) & \phi_1(x_N) & \cdots & \phi_{M-1}(x_N) \end{bmatrix}. \quad (14)$$

Finally, the maximum posterior probability with the given dataset D is defined by Eq. (15),

$$\arg_{\omega} \max P(\omega | D^M) = m_M. \quad (15)$$

Curve Fluctuation Fault. Curve fluctuation fault means that some abnormal fluctuations exist in the action stage or release stage. For this kind of fault, this paper utilizes harmonic superposition method for simulating curve fluctuation fault $F(x; t)$, which superposes at least one special triangular wave series on the $f^*(x; t)$,

$$F(x; t) = f^*(x; t) + \sum_{i=1}^N A_i \sum_{k=0}^{\infty} (-1)^k \frac{\sin((2k+1)t_k)}{(2k+1)^2} \quad (16)$$

where A_i and N are respectively the amplitude of the i -th wave and the total number of the wave.

Fixed Type Fault. Fixed type faults are generated by particular reasons and can be simulated by combining $f^*(x; t)$ and particular rules shown in Table 3.

Table 3. Fixed constraint faults and corresponding constraints

Fault type	Constraints	Remarks
Conversion current exceeded	$x_i \geq \theta, i \in [n_1, n_2]$	$\theta \in [1.2x_i^*, 1.3x_i^*]$
Electric relay 2DQJ turnout failure	$\begin{cases} x_i \equiv 0, i \in [1, 3n] \\ x_j \equiv x_2, \textit{otherwise} \end{cases}$	
Blocking in the gap	$x_i \equiv 0, i > n_2$	
Open start-up circuit	$x_i \equiv 0, i \in [0, 3n]$	$\begin{cases} \{\gamma_i \sim N(\mu_i, \sigma^2)\}_{i=1,2,3} \\ \mu_1 < \mu_2 < \mu_3 \end{cases}$
Release failure	$x_i = x_i + \varepsilon, i \in [\gamma_1, 3n]$	
Turnout blocked	$x_i = x_i + \varepsilon, i \in [\gamma_2, 3n]$	
Lock failure	$x_i = x_i + \varepsilon, i \in [\gamma_3, 3n]$	

4 Simulation Data Generation Algorithm Validation and Quality Evaluation

4.1 Simulation Data Generation Algorithm Implementation

In this experiment, a total of 1,931 current curves, including 15 faults were collected from Guangzhou railway bureau in China and the template curve of these data is shown in Fig. 2. The simulation results are shown in the following parts.

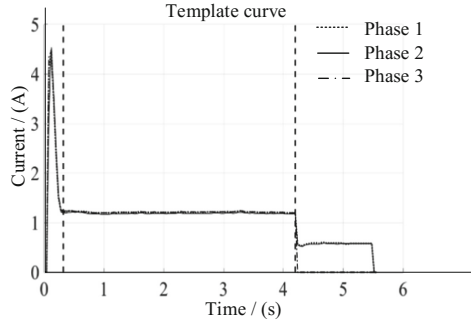


Fig. 2. Template curve

Time Anomaly Fault Simulation (M_1, M_2). Assuming that the variance of noise terms β^{-1} in Bayesian Regression is 0.05, the extra data length generated by regression is defined as the half of the corresponding initial length. The comparison between the simulation fault and the actual fault is as follows (take the first phase as an example) (Fig. 3):

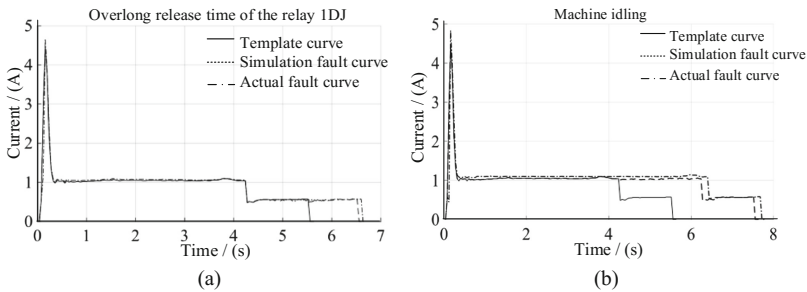


Fig. 3. Time anomaly simulation faults and corresponding actual faults

Curve Fluctuation Fault Simulation (M_3, M_4). In this paper, a triangular wave with an amplitude of 0.5 A and a period of 2π is used to generate the M_4 fault. For the fluctuations of current curves, seven triangular waves with an amplitude of 0.5A and a period of 2π are used. The generated fault is shown in Fig. 4.

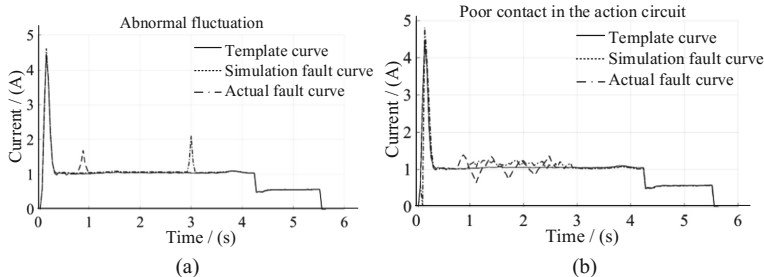


Fig. 4. Simulation of abnormal fluctuations and corresponding actual faults

Fixed Type Fault Simulation ($M_5 \sim M_{11}$). Suppose that the parameters of the medium exceed $\theta = 1.1x_7^*$; The indexes in M_7 is $i \sim N(0.5, 5e^{-4})$; The indexes in M_8 is $i \sim N(2, 5e^{-4})$; The indexes in M_9 is $i \sim N(4, 5e^{-4})$; The generated failure diagram and the corresponding actual failure diagram are shown in Fig. 5.

4.2 Quality Evaluation of Simulation Data

At present, many intelligent algorithms for fault diagnosis, diagnosis prognosis and early warning employ the distance method based on the maximum likelihood principle.

There are an extensive collection of particular distances have been put into use in different industry fields, such as Chebyshev distance and Minkowski distance. In this paper, we utilize the Derivative Dynamic Timing Warping (DDTW) [9] as the distance indicator. DDTW calculates the distance between two time series with different lengths by warping them into the same dimension. As a result, the smaller the DDTW value is, the closer in shape the two curves is, and definitely, the higher the similarity is. To quantify the performance of SFGA, this paper randomly chooses eleven simulation faults as candidate curves and the DDTW distance between the candidate dataset and corresponding real fault curves are shown in Table 4.

In Table 4, M_i and M_i^* represents the simulation fault and corresponding real fault of class i . According to DDTW values shown in Table 4, the following conclusions can be drawn:

- (1) Each type of simulation fault is the most similar to its corresponding real fault, indicating that the simulation faults generated by SFGA can be accurately mapped to the corresponding real faults, i.e., this paper can generate all kinds of faults of any speed-up turnout.
- (2) DDTW values only effectively represent the distance between two curves with different shapes but different durations. Therefore, the two types of time abnormal faults (M_1 and M_2) are similar to other types of real faults.
- (3) Due to the randomness of curve fluctuation in time and amplitude, Therefore, DDTW distance between M_3 and M_3^* can be relatively large.

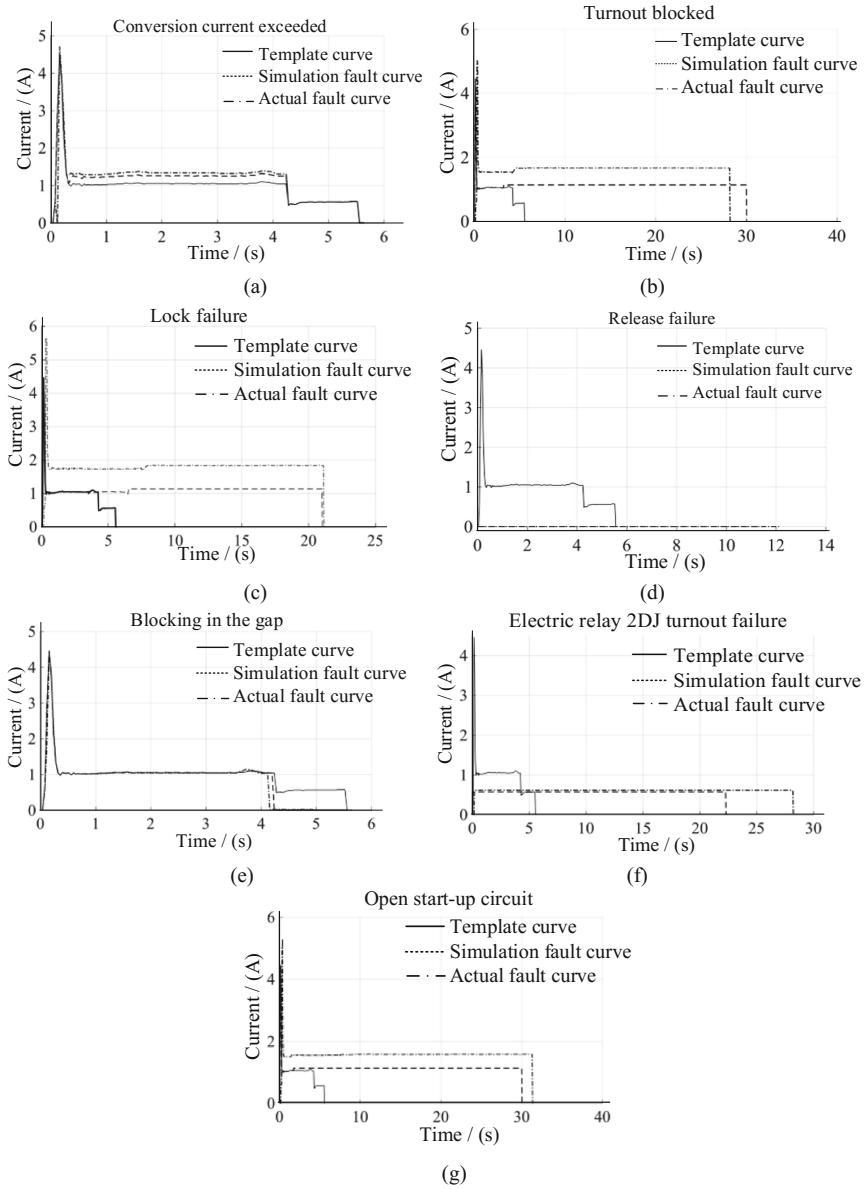


Fig. 5. Simulation of fixed constraint faults and corresponding actual faults

Table 4. DDTW between simulation fault and actual fault

	M_1^*	M_2^*	M_3^*	M_4^*	M_5^*	M_6^*	M_7^*	M_8^*	M_9^*	M_{10}^*	M_{11}^*	f^*
M_1	3.9	4.3	7.4	6.3	6.0	16.3	33.2	36.4	39.8	6.1	24.2	5.6
M_2	6.9	5.2	8.2	5.3	7.0	25.6	25.6	29.5	34.2	6.2	31.1	6.3
M_3	8.2	8.0	6.5	14.0	10.1	28.1	28.1	32.2	38.0	8.3	18.0	8.5
M_4	18.3	24.6	16.0	1.2	19.	87.1	87.1	96.2	95.7	33.1	85.8	22.9
M_5	6.1	6.2	4.0	19.1	3.5	93.0	93.0	93.4	95.6	13.2	76.3	8.3
M_6	16.2	18.7	20.2	19.0	18.2	86.1	86.1	83.7	76.1	20.2	44.5	16.8
M_7	37.4	37.4	35.4	26.2	27.2	1.3	83.7	8.3	22.2	35.9	96.8	23.5
M_8	40.5	40.5	38.8	29.6	86.1	8.3	76.1	3.2	15.4	34.6	92.4	28.1
M_9	38.3	39.3	38.7	32.7	83.7	22.2	20.3	15.4	4.3	28.9	98.1	33.8
M_{10}	6.2	5.1	8.3	10.5	76.1	35.9	44.5	34.6	28.9	99.1	29.3	5.9
M_{11}	23.4	27.3	25.7	24.8	24.4	44.5	99.8	88.4	98.1	29.3	0.0	20.7

5 Conclusion

In this paper, aiming at tackling the limited and incomplete fault data problem of speed-up turnouts, the SFGA is proposed and can generate large simulation fault data of speed-up turnout. It employs an extraction method for different turnouts and creates an adaptive segmentation-based template curve. SFGA also offers a simulation data generation model to combine particular rules and template curves and produce simulation faults. Furthermore, the faults generated by SFGA can be used as an extra set of real faults and also can be used for certifying fault diagnosis algorithms of speed-up turnouts, which lays a solid foundation for the further study, especially for quick fault warning and diagnosis of speed-up turnouts. Future work will strive to apply SFGA for undefined turnout faults to pursue a broader applicability.

Acknowledgement. The project is supported by the National Key R&D Program of China (2016YFB1200401).


References

1. Zhou, F., Xia, L., Dong, W., et al.: Fault diagnosis of high-speed railway turnout based on support vector machine. In: IEEE International Conference on Industrial Technology. IEEE (2016)
2. Eker, O.F., Camci, F., Guclu, A., et al.: A simple state-based prognostic model for railway turnout systems. IEEE Trans. Industr. Electron. **58**(5), 1718–1726 (2011)
3. Lu, E.B.: Signal Centralized Monitoring Information Analysis Guide. China Railway Press, Beijing (2015)
4. Guo, Z., Ye, H., Dong, W., et al.: A hybrid feature extraction method for fault detection of turnouts. In: Chinese Automation Congress (CAC), pp. 540–545. IEEE (2017)
5. He, Y.M.: Research on Fault Diagnosis Method of High-Speed Railway Turnouts. Beijing Jiaotong University, Beijing (2014)

6. Wang, R.F., Chen, W.B.: Research on fault diagnosis method for S700K switch machine based on grey neural network. *J. China Railw. Soc.* **38**(6), 68–72 (2016)
7. Gullo, F., Ponti, G., Tagarelli, A.: Clustering uncertain data via k-medoids. *International Conference on Scalable Uncertainty Management*, pp. 229–242. Springer, Heidelberg (2008)
8. Minka, T.: Bayesian linear regression. Technical report. MIT (2000)
9. Keogh, E.J., Pazzani, M.J.: Derivative dynamic time warping. In: *Proceedings of the 2001 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, pp. 1–11 (2001)



Research on the Warning Threshold of Rail Transit Passenger Flow by Big Data

Tengfei Yuan^(✉), Xiaoqing Zeng, Qipeng Xiong,
and Chaoyang Wu

The Key Laboratory of Road and Traffic Engineering,
Ministry of Education, School of Transportation Engineering,
Tongji University, 4800 Cao'an Road, Shanghai, China
{15yuantengfei, zengxq, wcy}@tongji.edu.cn,
269365547@qq.com

Abstract. Recently there are a lot of crowd-related accidents (panic/crowd/stampede) occurring in rail transit hub, which is mainly caused by the large passenger flow. Due to the research of rail transit passenger flow warning is lack of practical value, and the relevant standards cannot be adapted to the local conditions. Therefore, this research adopts the mobile phone signaling, Wi-Fi detection technology and video image processing technology to obtain the accurate passenger flow data. Based on the research of passenger flow warning indicators, the passenger flow density is regarded as the warning indicators, so the precise region area becomes essential. Then the spatial map technology is used to divide the regions of rail transit hub and calculate the region area by the spatial geometric cross processing method. On this foundation, this paper constructs the warning thresholds of rail transit passenger flow which mainly includes the absolute warning threshold, relative warning threshold and abrupt warning threshold. Finally, the warning thresholds are verified in the Hongqiao Rail Transit Hub, which can warn the large passenger flow timely and precisely. In addition, the warning threshold not only can satisfy the demand of the passenger flow management, but also can used in different conditions.

Keywords: Rail transit · Passenger flow · Warning threshold · Big data · Spatial map technology

1 Introduction

Rail transit plays an important role in urban transportation system, which has become the main means to solve the serious traffic problems, because its advantages of large traffic volume, fast speed, punctuality, safety etc. [1]. In this situation, the passenger flow and mileage of urban rail transit is increasing sharply day by day, in China. Recently the statistical data has shown that there are 14 lines, 366 stations and 617 km operation mileage in Shanghai by the end of 2015. Besides, the average daily passenger flow of Shanghai rail transit was about 8395 thousands people in 2015, even the maximum is up to 10360 thousands during the holidays [2]. What is more, the passenger flow of rail transit

hub is especially large, so the crowd environment is easy to cause the passenger panic, malignant stampede and evacuation difficulty etc. [3–5].

Owing the passenger flow volume is far beyond anticipation, the Shenzhen Rail Transit Line 1 was suspended for 42 min, on 2005 New Year's day [5]. Even the stampede accident occurred in Belarus subway station due to the excessive density of passenger flow, which resulted 54 deaths and a lot of casualties, in 1999 [7]. Guangzhou Rail Transit Line 5 occurred the stampede accident, and caused 13 people were injured, in March 2014. Shenzhen Rail Transit occurred the stampede incident, resulted 12 people were injured [8]. Therefore, it is necessary to warn the rail transit passenger flow, so as to ensure the safety of travelers. But the research of rail transit passenger flow warning is usually limited to the experience analysis, or study the large passenger flow features in some specific scenarios, such as Shanghai World Expo [9], Xi'an Metro [10] and Beijing South Railway Station [11]. Therefore, the practicability and accuracy of them are difficult to guarantee.

In order to remedy these deficiencies, the paper regards the warning threshold of rail transit passenger flow as research subject, which is essential to warn the passenger flow. Firstly, this research adopts the advanced big data technology and data fusion method to collect the accurate passenger data. Secondly, the passenger flow density is selected as the warning indicator based on the advanced spacial map technology which can calculate the precise region area. Thirdly, according to analyze the features of urban rail transit, the warning thresholds of rail transit passenger flow are illustrated. Eventually, the warning thresholds are verified in the Hongqiao Rail Transit Hub, which can warn the large passenger flow timely and precisely.

2 Data Collection and Fusion

2.1 Data Collection Method

At present, the dominated technologies of passenger flow data collection contain the mobile phone signaling, Wi-Fi detection technology and video image processing technology, etc., which process different characteristics in passenger flow data collection (12).

2.2 Data Fusion Method

Mobile signaling, Wi-Fi detection and video image processing technology are applied in corresponding region of rail transit hub, which can effectually address the passenger detection of the indoor passenger flow at horizontal and vertical level. And according to the data fusion, this research can obtain the accurate passenger flow data, so the research of warning threshold processes the solid foundation, which can improve the emergency capacity of rail transit hub administration department. The data fusion model and data processing flow are shown as Fig. 1.

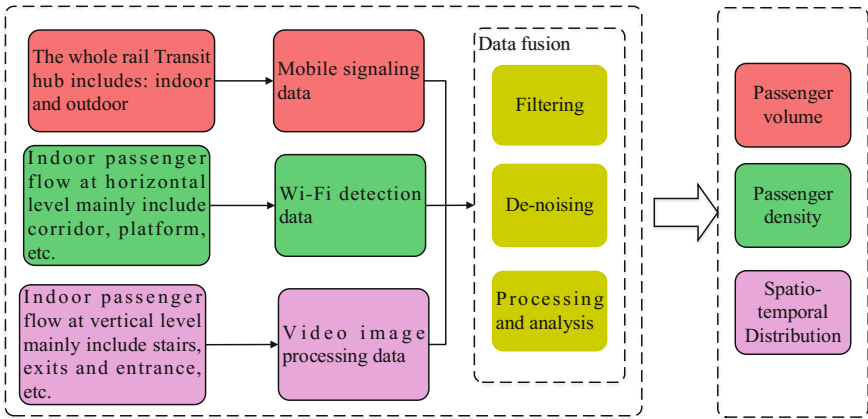


Fig. 1. Data fusion model and data processing flow.

3 Research of Passenger Flow Warning Indicators

3.1 Selection of Passenger Flow Warning Indicators

In previous studies, passenger flow warning correlation indicators mainly include: service intensity, passenger space load, passenger flow density and speed, duration of heavy passenger flow congestion and the number of passengers affected, which can well reflect the status of passenger flow, but it is difficult to obtain [12].

With the development of advanced technologies, this research can directly obtain the passenger flow density of some area by using the big data and map technology. So the passenger flow density can be realized as the warning threshold, which not only can be used as an indicator of passenger flow classification, but also can describe the status of passenger flow.

3.2 Calculation of Passenger Flow Density

Passenger Flow Density Calculation Process. First of all, select the monitoring region by the data-based map and calculate the area of the selected. Secondly, compute the real area of selected region according to the scale coefficient. Finally, use the fore-mentioned method to count the population number according to the principle of point set topology and the method of spatial superposition analysis, shown in Eqs. (1) and (2).

$$\rho = \alpha \frac{N}{S_{map}} \tag{1}$$

$$\alpha = \frac{1}{r^2} \tag{2}$$

Where

- N : The passenger flow volume by data fusion;
 S_{map} : The area of selected on the map based on the geometric cross and curve discretization method;
 α : The scale coefficient;
 r : The scale of map;
 ρ : The passenger flow density of the selected region.

Area Calculation on the Map. This section introduces the area calculation of the selected region on the map, which is the key part of calculating the passenger flow destiny accurately.

The research uses of the Baidu map as the base map, so the users can select any region according to the latitude and longitude translated by location data. First of all, transform the vector map into grid map by coordinate transform. Then calculate the area of selected region on the map by spatial geometric cross processing method. That is drawing the vertical lines of the X axis along each vertexes of a polygon, and calculates the area which is surrounded by each sides, two vertical lines and the X-axis. For a polygon with a hole or an inner island, the area equals to the difference between the outer polygon area and Inner Island [13], as described in Eq. (3).

$$S = \sum_{i=1}^N (S_{vi} - S_{ui}) \quad (3)$$

Where

- S_{vi} : The area of effective region which is enclosed by the curve and the X-axis;
 S_{ui} : The area of effective region which is enclosed by the other curve and the X-axis.

In addition, this method is also suitable for different area calculation, the concrete formula is shown in Eq. (4).

$$S_{map} = \frac{1}{2} \left[\sum_{i=1}^{N-2} (x_i y_{i+1} - x_{i+1} y_i) + (x_N y_1 - x_1 y_N) \right] \quad (4)$$

Where

- N : The number of polygon vertices;
 y_i : The plane coordinate of the corresponding vertex.

4 Warning Threshold of Rail Transit Passenger Flow

After analyzing the features of the rail transit passenger flow, this research proposes the comprehensive warning thresholds based on the calculated passenger flow density, which mainly include the absolute warning threshold, relative warning threshold and abrupt warning threshold. By establishing the floor plan of Hongqiao hub to study the warning threshold of rail transit passenger flow, which is presented in Fig. 2.

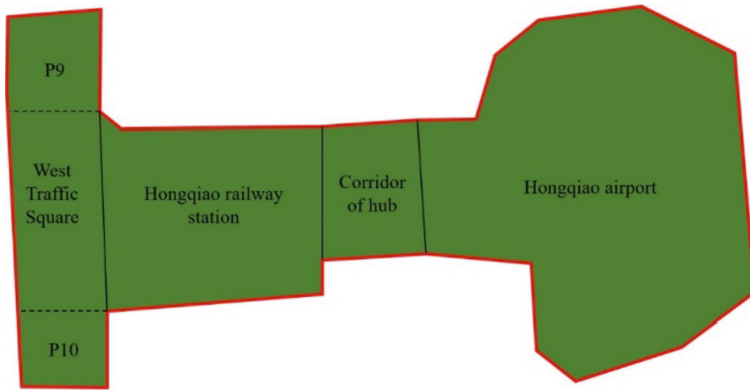


Fig. 2. Shanghai hongqiao hub.

4.1 Absolute Warning Threshold

Absolute threshold which is mainly determined by the current popular international service standard of the traffic facilities, and it can be calculated through computing the capacity of various facilities. Owing to the absolute threshold is related to the properties of the service facility, so it will not change with time.

The absolute threshold mainly refers to the current international service standard of footpath and queuing zone, which main contains the service level, per capita area, flow rate velocity and V/C, as shown in Tables 1 and 2 [14].

Table 1. Standard of service level for the footpath zone

Service level	Spatial (m^2 /person)	Flow rat (person/min/min)	Velocity (m/s)	V/C
A	>5.6	≤ 16	>1.30	≤ 0.21
B	>3.7–5.6	>16–23	>1.27–1.30	>0.21–0.31
C	>2.2–3.7	>23–33	>1.22–1.27	>0.31–0.44
D	>1.4–2.2	>33–49	>1.14–1.22	>0.44–0.65
E	>0.75–1.4	>49–75	>0.75–1.14	>0.65–1.0
F	≤ 0.75	Varietal	≤ 0.75	Variable

Table 2. Standard of service level for the queuing zone

Service level	Spatial (m ² /person)
A	>1.2
B	>0.9–1.2
C	>0.3–0.6
D	>0.6–0.9
E	>0.2–0.3
F	≤ 0.2

Accordinging the standard of service level for the footpath zone and queuing zone to determine the absolute threshold of each sections of Hongqiao hub, as presented in Table 3.

Table 3. The absolute threshold of hongqiao hub

Level	Density (m ² /per)
I	5.25
II	3.47
III	2.02
IV	1.16
V	0.59

4.2 Relative Warning Threshold

Relative threshold which is determined by the historical data and it can pick out the abnormal data of passenger density according to the normal distribution model. For the actual case of this study, the anomaly data means the large passenger flow occurs. Due to the threshold changes over time and condition, so it is called the relative threshold.

On the basis of normal distribution, it can be known from the Statistics, as presented in Eq. (5):

$$\begin{aligned}
 P\{\mu - \sigma < x < \mu + \sigma\} &= 0.6826 \\
 P\{\mu - 2\sigma < x < \mu + 2\sigma\} &= 0.9544 \\
 P\{\mu - 3\sigma < x < \mu + 3\sigma\} &= 99.74\%
 \end{aligned}
 \tag{5}$$

Where

- P : The probability;
- μ : The mean value;
- σ : The standard deviation.

It also means that more than 99% of the data can be in the normal range, which is called 3σ theorem.

According to the 3σ theorem to set up the appropriate threshold, which is flexible and can be adjusted by the time and practical condition. The relative threshold of Hongqiao each section is presented in Table 4.

Table 4. The relative threshold of hongqiao each section

Relative threshold	West Traffic Square	Hongqiao railway station	Corridor of hub	Hongqiao airport
Workdays	8481	14722	175	7368
Weekends	10241	17570	335	9367

4.3 Abrupt Warning Threshold

Abrupt threshold mainly aim at the abrupt peak and random fluctuation phenomenon. On the foundation of the relative threshold, this research adopts the abrupt threshold to improve the sensitivity of changes in perceived passenger flow.

The abrupt threshold is fit for the following conditions:

- (1) The region that it usually occur the extreme passenger flow phenomenon, such the corridors and platform of rail transit hub;
- (2) The Queuing zone that it is easy to swarm a large number of population. Therefore, it is necessary to monitor these zones with the abrupt threshold.

Due to the absence of uniform standard of abrupt threshold, it is necessary to define it. Considering the characteristics of passenger flow, the slope of line can directly represent the variation of the passenger flow, as shown in Fig. 3, the slope between point A and point B is $(30 - 10)/\Delta t$, the slope between points C and D is $(70 - 50)/\Delta t$. This slope is equal to the variance ratio of passenger flow, so it is can be used to measure the relative degree of mutation.

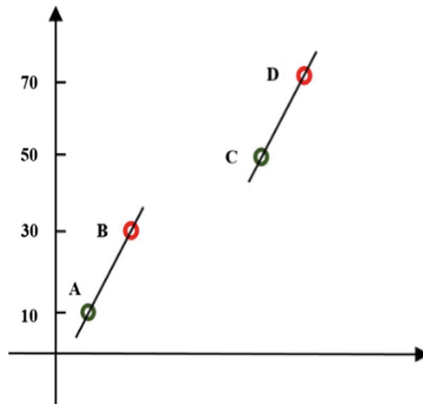


Fig. 3. Abrupt threshold diagram.

4.4 Test and Verify

After establish the warning thresholds, this research test and verify them in Shanghai Hongqiao rail transit hub, which is comprehensive and significant station. Due the large passenger flow phenomenon usually occurs, so it is fully necessary to establish the appropriate warning thresholds for it.

Data Accuracy Test. This research adopts the counted manually, and compare with the result of data fusion. Take the West Traffic Square as an example, and this region has been divided into A–J, 10 subsections as shown in Fig. 4.

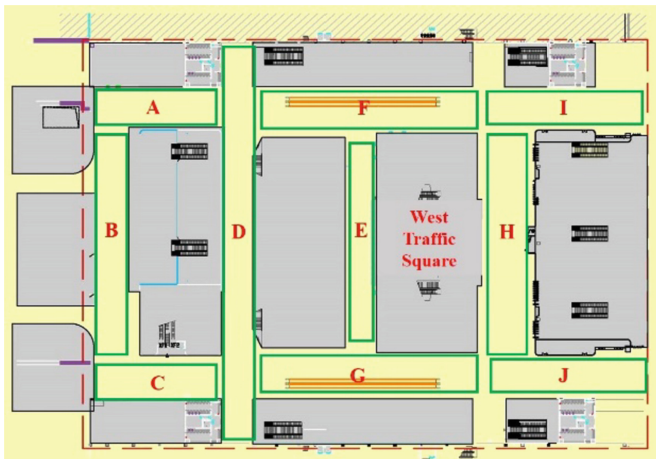


Fig. 4. The West Traffic Square and its 10 subsections.

The analysis result is presented in Table 5, and the result shows that the data accuracy satisfies the research demand.

Table 5. The result of data accuracy test

Subsection	Counted manually	Data fusion result	%
A	38	33	86.84
B	15	18	80.00
C	12	10	83.33
E	8	5	62.50
F	5	4	80.00
G	56	54	96.43
H	78	79	98.72
I	98	102	95.92
J	39	43	89.74

Area Calculation Test. According to the area calculation on the map, the area of West Traffic Square is calculated, and compares with the manual measurement result, which is shown as Table 6, and area calculation accuracy also satisfies the demand of research.

Table 6. The result of area calculation test

Zone (m ²)	Calculated area	Measurement	%
West Traffic Square	9177	9005	98.13
Hongqiao railway station	28500	27500	96.49
Corridor of hub	82178	80015	97.35
Hongqiao airport	9883	9506	96.12

Warning Threshold Verification. Ultimately, this research applies the warning threshold to verify in the July 2017. When the summer vacation is coming, there are lots of passengers in Hongqiao rail transit hub. The paper uses the alarm numbers, miss numbers and accuracy rate to verify the passenger flow warning threshold. The dialed results are listed in Table 7, and it is can be concluded that the rail passenger flow warning threshold not only can accurately warn the lager passenger flow, but also can fully satisfy the demand of the management department of Hongqiao Hub.

Table 7. The results of warning threshold verification

Zone	Alarm numbers	Miss numbers	Accuracy rate
West Traffic Square	9	0	100%
Hongqiao railway station	21	1	95.24%
Corridor of hub	8	0	100%
Hongqiao airport	12	0	100%

5 Conclusion

In this paper, we adopt mobile signaling, Wi-Fi detection and video image processing technology to collect the passenger flow data, as well as uses the spatial map technology to calculate the area of rail transit hub timely and accurately. On this foundation, this research selects the passenger flow density as measure indicator according to the practical situation, and constructs the warning threshold of rail transit passenger flow which includes the absolute threshold, relative threshold and abrupt threshold. And Tested and verified in the Hongqiao Hub, the results is acceptable. Therefore, this research can guarantee the practicality and reliability of the passenger flow warning, as well improve the processing capacity of the passenger flow management department of Hongqiao Hub.

Acknowledgement. The project is supported by Tongji University and the project “Research on the Practice and Improvement of the Construction Innovation Technology of Shanghai Rail Transit Line 17” (Number JS-KY18R022-7). The authors are grateful for the reviewer of initial drafts for their helpful comments and suggestions.

References

1. Boyce, D.: Urban transit: operations, planning, and economics, ed. by V.R. Vuchic. *J. Reg. Sci.* **46**(3), 566–568 (2010)
2. Shao, J., Shi, W.: The rights model of disabilities and shanghai rail transit - current wheelchair accessibility in shanghai rail transit. In: *TRANSED 2010: 12th International Conference on Mobility and Transport for Elderly and Disabled Persons* (2010)
3. An, S.: Inspiration for rail transit passenger flow forecast from the operation of beijing subway line 5. *Urban Rapid Rail Transit* (2008)
4. Ma, C.Q., Wang, Y.P., Guo, Y.Y., et al.: Sensitivity analysis on urban rail transit passenger flow forecast. In: *International Conference on Electric Technology and Civil Engineering*, pp. 1537–1541. IEEE (2011)
5. Zhou, Z., Zhang, Y.: Influence of bus fare adjustment on rail transit passenger flow. *Urban Public Transp.* (2009)
6. Zhang, J., Zhang, M., Wang, K.: Study on vehicle maintenance technology and equipment of suspension rail transit. *Mod. Urban Transit* (2016)
7. CBS NEWS. <https://www.cbsnews.com/news/54-crushed-to-death-in-belarus/>. Accessed 31 May 1999
8. Ming-Wei, H.U.: A survey and simulation of passenger flow organization of the Shenzhen urban rail transit station. In: *Cota International Conference of Transportation Professionals*, pp. 2991–2997 (2011)
9. Li, Z.: 2010 Shanghai world expo passenger flow balance and guidance under background of information. *Public Utilities* (2010)
10. He, C.L.: Passenger flow forecast based on grey markov model—a case of Xian metro liane 2. *Archit. knowl.* (2016)
11. Zhang, S.F., Zhang, X.: Analysis on dynamic simulation for passenger flow line organizations and optimizations in Beijing south railway station. *Railw. Comput. Appl.* (2010)
12. Daamen, W.: Modeling passenger flows in public transport facilities. *Free Flow Speeds* (2004)
13. Hong-Xia, Q.I., Zhang, M.S., Tang, Z.Q., et al.: Vector quantization based on wavelet transform grid topographic map automatic optimization research. *Geomat. Spat. Inf. Technol.* (2015)
14. Highway Capacity Manual (2010). <http://hcm.trb.org/?qr=1>. Accessed 2010



Optimization of Subway Departure Timetable by Genetic Algorithm

Junxiang He, Xiaoqing Zeng, Peiran Ying, Xinchun Xu,
and Yizeng Wang^(✉)

The Key Laboratory of Road and Traffic Engineering, Ministry of Education,
Tongji University, No. 4800 Cao'an Road, Shanghai 201804, China
1102790744@qq.com

Abstract. This paper the objective constraints such as the number of train groups available for use and the basic interests of the operating company, and optimizes the departure time of urban rail transit according to the existing passenger flow characteristics change. Based on the characteristics of passengers' travel behavior and the habits of taking urban rail transit, a mathematical model with the objective function as the average waiting time of passengers is constructed. Using genetic algorithm to solve the model reasonably, it is found that the optimized scheme can effectively shorten the average waiting time of passengers under the conditions of satisfying various constraints.

Keywords: Genetic algorithm · Urban rail line · Timetable planning · Matlab

1 Introduction

Zhou and Zhong (2005) put a random accident scenario into a train schedule optimization model to optimize both the expected waiting time of high-speed train passengers and the total running time of high-speed and medium-speed trains. Goverde (2007) describes a method of using max-plus software systems to measure the stability of railway schedules and analyze the spread effects of railway delays. Liebchen (2008) used a script that lets emergencies occur periodically to optimize the timetable of the Berlin Metro to reduce passenger waiting time. Wong (2008) focuses on how to simultaneously optimize the entire transport network to minimize passenger waiting time.

2 Optimization Model of Departure Timetable

2.1 Problem Analysis

2.1.1 Characteristics of Subway Train Operation

Suppose a subway line has a total of N stations in one direction, then there will be $2n$ stations in both directions. In this paper, it is stipulated that the direction of the subway from the departure station is positive and the sequence of arrival codes is $1, 2, 3 \dots n$. After arriving at the forward terminal, the turning head is turned to run again. At this time, the driving direction is set in the opposite direction and is encoded as

$n + 1, n + 2, n + 3 \dots 2n$ in sequence according to the arrival order. Once again back to the departure station, the train will enter the preparation time for the next drive.

2.1.2 Characteristics of Passenger

Compared with intercity railway, such as high-speed rail, passengers pay less attention to subway timetable when they take subway and other urban rail transit. They go directly to the platform to wait for the first train. If they can not go, they wait for the next train. The trip behavior will not be changed by the adjustment of departure time plan.

2.1.3 Hypothesis of Waiting Time for Passengers at Peak Hours

This paper assumes that passengers can all get on the train to simplify the optimization and choose the passenger flow as far as possible when choosing the example, compared with the maximum capacity of the train is not easy to occur because of too many passengers, the train can not let the passengers waiting on the platform get on the train at once.

2.1.4 Reasonable Arrangement of Trains and Routes

It is particularly important to coordinate and utilize the train and track resources reasonably. Otherwise, it is easy to appear the situation of “wired without train” and “wired without train”. For the general rail passenger transport, the whole flight arrangement is more systematic and passengers buy tickets in advance and travel on time, so it is not easy to waste resources or no resources can be provided. In the railway track freight transportation, the dispatching department of our country often adopts the method of “losing line” when dealing with the situation of “wired and no flow”.

2.2 Mathematical Model

Based on the analysis of the characteristics of the known passenger flow in Nanjing, the operation characteristics of the subway are analyzed, the parameters are selected and the model is established to solve the optimal departure interval of the subway timetable.

2.2.1 Basic Parameters

The parameters that will be used in the modeling process and their corresponding symbols are as follows:

1. the total number of subway g

Represents the number of metro vehicles set up on a predefined optimized subway line.

2. the number of run k and the total number of run m

K denotes the number of run of trains in the planned period, and M denotes the total number of run of trains in the planned period.

3. time t
4. the number of the stations i

Each platform will be numbered separately according to the direction of the train and the position in the whole line. I numbers range from 1, 2, 3, ... 2n, and remember that Station I and Station 2n + 1-I are numbers of the same platform in different directions.

5. Passenger arrival rate pas (i,j,t)

Passenger arrival rate PAS (i, j, t) of station J is the destination of station I at time t. The unit is person/min. At the same time, it is clear that all passengers will alight when the train arrives at the forward terminal, so no passengers will ride from the forward terminal to the reverse, that is, the arrival rate of passengers at the right and the right time pAs (I, J, t) is always zero.

6. the time for train No. K to arrive at I station $A_{i,k}$

It means that the time for train No. K to arrive at I station is accurate to minutes.

7. the time that the subway comes from station I to station I + 1 $B_{i, i+1}$

Indicates the time required for the subway from station I to station I + 1, including the stopping time at station I and the running time required for leaving station I to station I + 1. For the convenience of calculation, the accuracy is taken to minutes.

8. the time for train No. K to departure I station $D_{i,k}$

Because all the time units in this optimization scheme are all value to minutes, and the subway train usually stops at one station for less than one minute, so the default train departure time $D_{i,k}$ is equal to the train arrival time $A_{i,k}$.

9. $P_{i,k}$ and $Z_{i,k}$

When the train No. K leaves platform No. I, the number of passengers on board is $P_{i,k}$ and the carrying rate of passengers is $Z_{i,k}$.

10. Minimum preparation time B_0

When each train reaches Platform 2n, it needs to enter the platform to rest at least one minimum preparation time B_0 before it can start again.

11. Remainder $L_{i,k}$

The number of passengers arriving at the destination after the arrival of the k th train at station I has not yet reached the number of passengers boarding the station

12. Maximum carrying capacity of train PN

In theory, the maximum passenger carrying capacity of a train is equal to the maximum passenger carrying capacity corresponding to the type of single carriage used by the train group (calculated by 9 persons per square) multiplied by the number of cars connected by each train group.

2.2.2 When the Train Leaves the Station, the Number of People on the Train $P_{i,k}$

When train No. K arrives at I station, it is assumed that the passengers should obey the principle of “first down and back up”. Then $P_{i,k}$ it can be obtained through distributed computation.

$$L_{i,k} = P_{i-1,k} - \sum_{j=1}^{i-1} \sum_{t=A_{j-1,k-1}+1}^{A_{j-1,k}} pas(j, i, t) \quad k = 1, 2, 3, \dots, m; i = 2, 3, 4, \dots, 2n \quad (1)$$

$$P_{i,k} = L_{i,k} + \sum_{j=i+1}^{2n} \sum_{t=A_{i,k-1}+1}^{A_{i,k}} pas(i, j, t) \quad k = 1, 2, 3, \dots, m; i = 1, 2, 3, \dots, 2n - 1 \quad (2)$$

(1) is used to calculate the remaining number of passengers after getting off the train., (2) is used to calculate the number of people on the train leaving the platform.

2.2.3 Arrival Time of Trains

Because it is assumed in this paper that the departure time of the train is the same as the arrival time of the train at each station, the arrival time of the k th train at Platform I can be obtained from the arrival time of the train at the departure station plus the travel time of the stations on the way.

$$A_{i,k} = A_{1,k} + \sum_{l=1}^{i-1} B_{l,l+1} \quad k = 1, 2, 3, \dots, m; i = 2, 3, 4, \dots, 2n \quad (3)$$

2.3 Constraint

2.3.1 Restrictions on Vehicle Group Quantity

Since only Group G subway is put into operation on a subway line, no matter what timing scheme is adopted, the starting time of train K ($k > g$) must always be greater than or equal to the starting time of train K-G and the minimum preparation time. Expression is expressed by mathematical formula.

$$A_{1,k} \geq A_{1,k-g} + \sum_{i=1}^{2n-1} B_{i,i+1} + B_0 \quad (4)$$

2.3.2 The Restriction of the Minimum Passenger Rate $\min Z$

Here the following criteria are drawn up: if the number of $Z_{i,k} < \min Z$ is greater than the allowable number of occurrences Y, the arrangement of the K train is considered unreasonable and needs to be reconsidered.

2.3.3 Minimum Time Interval Between Adjacent Shift Vehicles J

In the subway operation, the departure time interval of the two subway trains must meet a minimum time interval value J, so as to meet the safety requirements.

$$A_{1,k} - A_{1,k-1} \geq J \quad k = 2, 3, 4, \dots, m \tag{5}$$

2.4 Objective Function

Suppose that if a group of passenger who want to go to station j arrive at station i at time t and t is between $A_{i,k-1}$ and $A_{i,k}$, then they will take the k th subway to their destination and the waiting will be $A_{i,k} - t$. So our objective function can be presented as follow

$$\min \sum_{k=1}^m \sum_{i=1}^{2n-1} \sum_{j=2}^{2n} \sum_{t=A_{i,k-1}+1}^{A_{i,k}} (A_{i,k} - t)^* pas(i, j, t) / PN \tag{6}$$

3 Genetic Algorithm

Genetic algorithm is a stochastic global search and optimization method which imitates the evolution mechanism of natural organisms. Its essence is an efficient, parallel and global search method. It can automatically acquire, accumulate and learn the knowledge about search space in the process of search operation, and adaptively control the search process to obtain the optimal results.

4 Simulation

4.1 Solving

In this paper, the genetic algorithm is used to solve the model. Chromosomes are encoded in traditional binary system. Each gene position corresponds to every possible departure time in the optimization period. For this time, “1” will have a train to send out, and “0” will not have a train to send out. The optimal time interval is 5 h and the unit is min. There are 300 possible departure times, that is, the individual chromosome length is 300. When calculating individual fitness, the constraints are combined with the objective function, and the individuals who do not meet the constraints will add penalty value to the objective function to reduce their fitness. Some parameters in the algorithm are as follows: the minimum time interval of trains is 2 min, the maximum carrying capacity of trains is 1 440, the minimum carrying rate of trains is 6.5%, the number of low-load passengers is 3, the number of trains available for dispatching is 15, the population size is 50, the number of genetic iterations is 500, and the crossover probability is 0.70, respectively. The probability of element variation is 0.5.

4.2 Data

Nanjing Line S1 is an important Metro extension line located in the south of Nanjing. The length of Nanjing Line S1 is 35.8 km, including 18.2 km underground, 0.7 km transitional and 16.9 km viaduct. Because this paper calculates the flow of people in minutes and the train stops at one stop for no more than one minute, it is of little significance to set the train stopping time. The train departure time is the train arrival time. It can be seen from the above that the early peak of Nanjing passenger flow is at 8–9, so the study of optimization period selected 6–11 points (Tables 1 and 2).

Table 1. Station number of Nanjing line S1

The name of the station	Forward numbering	Reverse number
Nanjing south railway station	1	16
Cuiping mountain station	2	15
Hohai University·Fo Cheng West Road Station	3	14
Jiyin Avenue Station	4	13
Central Square Road Station	5	12
Xiangyu Road North Station	6	11
Xiangyu Road South Station	7	10
Lu Kou Airport Station	8	9

Table 2. The time consumed between stations in Nanjing S1 line

Station number	$B_{i, i+1}$	Station number	$B_{i, i+1}$	Station number	$B_{i, i+1}$	Station number	$B_{i, i+1}$
1	5	5	6	9	6	13	4
2	4	6	5	10	5	14	4
3	4	7	6	11	6	15	5
4	5	8	4	12	5	16	4

The Nanjing Metro’s current record of passenger information is shown in Table 3.

Table 3. Nanjing metro passenger information record form

Arrive time	The time of trade	Kind of trade	The number of trade machine	The number of station	The trade station
2016/7/9 6:23	2016/7/9 6:38	2	33911060	44	71

Therefore, according to the information provided in the table above, the flow of people between any two stations in Nanjing S1 line can be calculated. At the same time, the flow of people from Nanjing S1 line to other lines and from other lines to S1

line should also be calculated. After calculation and collation, the following figure of Nanjing Line S1 passenger flow in the optimization period of change chart (1 min) (Fig. 1).

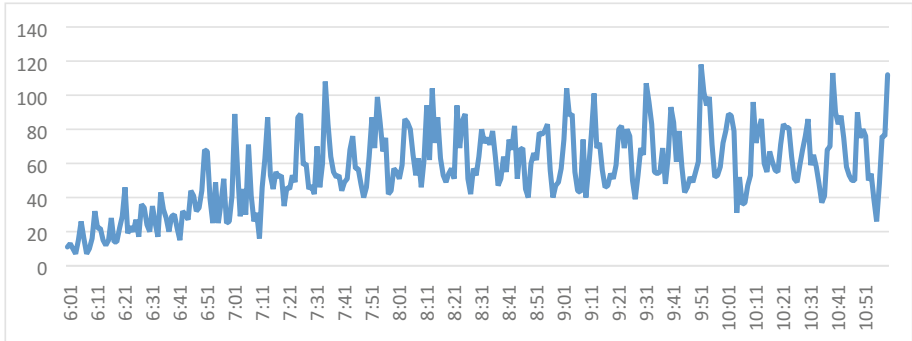


Fig. 1. Variation of passenger flow in Nanjing S1 line over time

As can be seen from the above chart, although passenger flow fluctuates greatly over time, the overall trend is always increasing and gradually stabilized.

4.3 Conclusion

Because of the randomness of the genetic algorithm, we have solved it many times and analyzed and compared the results of each calculation, so as to avoid getting unreasonable conclusions because of the premature convergence of the algorithm.

After solving the mathematical model and the genetic algorithm code, the following subway timetable is obtained from one of the optimal solutions. The optimal departure times are 46 times and the average waiting time is 4.98 min (Tables 4 and 5).

Table 4. The optimal departure time for all trains

Subway flight	Subway number	Departure time	Subway flight	Subway number	Departure time
1	1	6:15	24	9	8:55
2	2	6:28	25	10	9:03
3	3	6:40	26	11	9:06
4	4	6:50	27	12	9:12
5	5	7:01	28	13	9:16
6	6	7:09	29	14	9:23
7	7	7:15	30	15	9:29
8	8	7:23	31	1	9:34
9	9	7:28	32	2	9:42
10	10	7:37	33	3	9:49

Table 5. The optimal departure time for all trains

Subway flight	Subway number	Departure time	Subway flight	Subway number	Departure time
11	11	7:45	34	4	9:53
12	12	7:50	35	5	9:58
13	13	7:56	36	6	10:04
14	14	8:02	37	7	10:11
15	15	8:08	38	8	10:16
16	1	8:14	39	9	10:21
17	2	8:18	40	10	10:24
18	3	8:24	41	11	10:30
19	4	8:31	42	12	10:33
20	5	8:34	43	13	10:40
21	6	8:38	44	14	10:47
22	7	8:43	45	15	10:51
23	8	8:50	46	1	10:57

The train diagram of Nanjing Line S1 is drawn from the above plan. As shown in the following diagram, abscissa represents time, ordinate represents station, slant represents subway line, slope greater than 0 represents forward running, slope less than 0 represents reverse running, and slope equal to 0 represents preparation for running at the corresponding station (Fig. 2).

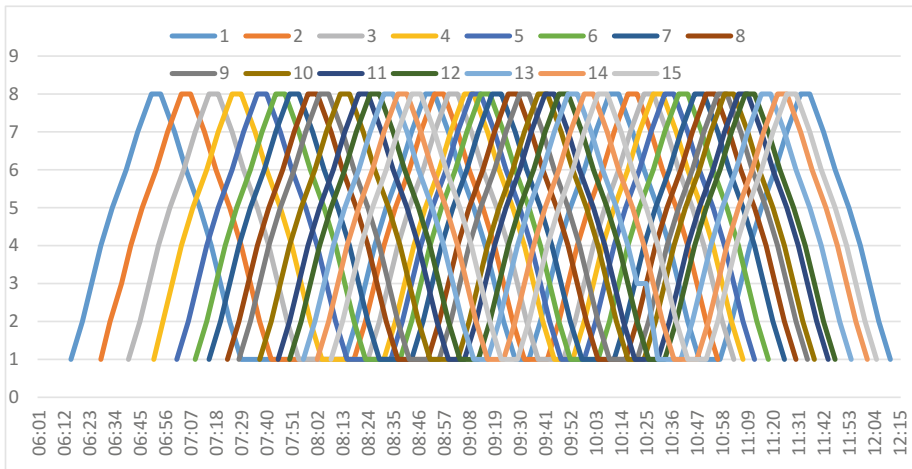


Fig. 2. Trains running from 6 a.m. to 11 a.m. on Nanjing line S1

Comparing the two charts with the previous passenger flow chart, it can be seen that the change of departure interval coincides with the change of passenger flow with the time interval. In reality, the train of Nanjing Line S1, which usually needs to wait for 10 min, can reduce the average waiting time to 4.98 min under the conditions of meeting the constraints of transferable train group and low-load passenger rate, greatly reducing the waiting time and providing great convenience.

5 Summary

The optimization results show that the waiting time of passengers will be significantly shortened after using the model. It proves that the model proposed in this paper has a certain role in solving the problem of average waiting time of passengers in urban rail transit.

References

1. Niu, H., Zhou, X.: Optimizing urban rail timetable under time-dependent demand and oversaturated conditions. *Transp. Res. Part C* **36**(11), 212–230 (2013)
2. Niu, H., Zhou, X., Gao, R.: Train scheduling for minimizing passenger waiting time with time-dependent demand and skip-stop patterns: nonlinear integer programming models with linear constraints. *Transp. Res. Part B* **2015**(76), 117–135 (2015)
3. Ceder, A.: Bus timetables with even passenger loads as opposed to even headways. *Transp. Res. Rec.* **1760**, 3–9 (2001)
4. Eberlein, X.J., Wilson, N.M., Bernstein, B.: The real time deadheading problem in transit operations control. *Transp. Res. Part B* **32**(2), 77–109 (1998)
5. Ceder, A.: Public-transport automated time tables using even headway and even passenger load concepts. In: *The 32nd Australasian Transport Research Forum*, pp. 1–17 (2009)



Review of Safety Assessment System on Urban Rail Transit

Songxue Gai^(✉) and Xiaoqing Zeng^(✉)

Key Laboratory of Road and Traffic Engineering of the Ministry of Education,
Tongji University, Shanghai 200092, China
{gaisongxue, zengxq}@tongji.edu.cn

Abstract. Safety is increasingly important for software based, critical systems of urban rail transit. This paper reviews and compares the advantages and disadvantages of various security assessment methods from qualitative, quantitative and comprehensive aspects. These methods are widely used in the field of practical safety assessment.

Keywords: Safety assessment · Urban rail transit · Railway risk assessment

1 Introduction

Urban rail transit is a complicated system, referring to a variety of work. With the increasing demand for urban rail transit, the requirements of subway reliability and security are getting higher and higher. Safety is central to rail operations [1]. Many methods are commonly used to analyze system safety, which breaks down system level hazards to failures of components, called (failure) events.

However, these methods are provided separately in each paper, there are few researchers systematically investigating the safety assessment of urban rail transport systems until now. This brings inconvenience to the selection of methods. Because we will be exhausted by integrate information from different papers. This paper reviews and compares various security assessment methods from qualitative, quantitative and comprehensive aspects. It will not only provide convenience for who want to do safety assessment on urban rail transit, and also a contribution to the field of safety assessment.

The paper is organized as follows: introduction, related concepts of safety assessment, safety assessment methods, and conclusion.

2 Concepts of Safety Assessment

Safety assessment refers to evaluate and analyze potential danger and its severity for a system with a specific function, and the quantitative index is given by the established index, grade or probability value, and finally the preventive or protective measures are decided according to the magnitude of the quantitative value. Related concept definitions are given in the Table 1 below.

Table 1. Concepts of safety assessment.

Concept	Definition
Risk	Risk is the potential of gaining or losing something of value
Hazard	Hazard is what causes harm or damage to humans, property, or the environment
Safety	Safety is the state of being “safe”, the condition of being protected from harm or other non-desirable outcomes. Safety can also refer to the control of recognized hazards in order to achieve an acceptable level of risk
Risk assessment	Risk assessment is the determination of quantitative or qualitative estimate of risk related to a well-defined situation and a recognized hazard [2]
Risk management	Risk management is the identification, assessment, and prioritization of risks followed by coordinated and economical application of resources to minimize, monitor, and control the probability and/or impact of unfortunate events or to maximize the realization of opportunities

3 Safety Assessment Methods

3.1 Qualitative Safety Assessment Methods

The qualitative safety assessment method is mainly based on the experience and intuitive judgment for qualitative ability on the urban railway system of technology, equipment, facilities, environment, personnel management and other aspects of the situation analysis, the results of the safety assessment are qualitative indicators, such as whether a safety index, accident category and factors causing the accident have been reached.

Typical qualitative safety assessment methods are expert evaluation, Safety Checklist Analysis.

Expert Investigation Method

Expert investigation method collects experts’ wisdom to make decision, avoiding subjectivity and unilateralism brought by individual’s views. There are two kinds of expert evaluation methods:

Delphi Method

Delphi method was first invented by Rand Corporation in 1964, and it is a group decision making behavior with anonymity, feedback and statistical characteristics, in essence, which are built on the expertise and experience of many experts and subjective judgment ability. Therefore, it is especially suitable for information analysis and prediction which is lack of information and historical data and is greatly influenced by other factors. Through a process of interaction with many experts, Delphi method makes the dispersed opinions converge gradually on the consistent results, and gives full play to the role of information feedback and information control [3].

Brain-Storming Method

The brain-storming method requires two meetings. The first meeting is a direct discussion of specific issues by experts; the second meeting is questioned by experts about the ideas presented at the first meeting. The main tasks are as follows: Study and discuss the problems that hinder the realization of imagination; demonstrate the feasibility of the proposed idea; discuss the limiting factors of the assumption and put forward some suggestions to eliminate the restrictive factors; In the course of questioning, the new constructive assumptions that arise are discussed.

The steps of expert investigation method shall be followed by the procedure:

- Make clear the concrete analysis and the forecast question;
- Form an expert review, analysis and forecasting team, which should be composed of forecasting experts, experts in the field of expertise, and deductive reasoning experts;
- Hold expert meetings to analyze, discuss and predict issues raised;
- Analyze and summarize the results of the expert meeting.

For safety evaluation, expert evaluation method is simple, objective, invited experts in the professional theoretical attainments and rich experience, and experts because of professional, security, evaluation, logic in the method of using logic reasoning and expert opinions are comprehensive, so the conclusion is comprehensive and correct. However, because of the high level of experts required to participate in the evaluation, not all projects are applicable to this method [4].

The expert evaluation method is applicable to the safety evaluation of analog engineering projects, systems and installations. The expert evaluation method is often used in the urban railway transit security assessment, and the evaluation method can be used to study and discuss the problem more deeply and thoroughly, and draw concrete executive opinions and conclusions, so as to facilitate scientific decision-making.

Safety Checklist Analysis

The safety checklist analysis is a commonly used method to analyze the railway transportation security system, which is a series of projects checklist analysis, to determine whether the state of the system and place meet the safety requirements, and through inspection, find hidden dangers in the system and improving measures. Inspection items can include the site, the surrounding environment, facilities, equipment, operation, management and other aspects [5].

The basis of compilation is established by relevant standards, procedures, specifications and regulations, accident cases and industry experience, and research results. The utility model has not only the advantage of simple and rapid analysis, but also the advantages of deep level analysis. It is an effective, simple and feasible method for safety inspection and potential danger detection. However, a large number of checklists must be prepared in advance for a variety of needs, and the quality of the safety checklist is generally limited by the knowledge level and experience of the staff members.

Failure Mode and Effect Analysis (FMEA)

In the early 50s, the United States for the first time applied FMEA thought to the design and analysis of a fighter operating system. Nowadays, different countries have different versions (such as Japan, Germany, the United States, etc.), the commonly used version in the United States is divided into DFMEA and PFMEA (the standard version of America's three largest car company). SFMEA is a German QMC-VDA version, including product system and process system.

Failure mode and effects analysis (FMEA) is a design analysis procedure involving the investigation and assessment of the effects of all possible failure modes on a system [6]. It is actually a combination of FMA (failure mode analysis) and FEA (failure effect analysis), and is failure modes and effects analysis or potential failure modes and effects analysis. As a tool for planning preventive measures, FMEA's main purpose is to find and evaluate the consequences of potential failure of product/process, to take measures to avoid or reduce potential failure and continue to improve. It can be easy and low-cost to modify a product or process, which can reduce the modification afterwards. So, ways can be found to avoid or reduce these potential failures. Because the product failure may be related to the design, manufacturing process, use, contractor/supplier, and service, the FMEA is subdivided into: DFMEA (Design FMEA), PFMEA (Process FMEA), EFMEA (Equipment FMEA), SFMEA (System FMEA). Among them, DFMEA and PFMEA are the most commonly used. The aims of this method include: Identify and evaluate potential failures in the product (process) and the consequences of the failure; Identify measures to eliminate or reduce potential failure opportunities; File the whole process.

FMEA can be a detailed analysis of a possible failure mode or an abnormal mode of operation for each component within a system, and infer its impact on the whole system, the possible consequences, and how to avoid or reduce the loss. The disadvantage is that we can only consider non-hazardous failure, and this always takes a long time, so we cannot consider all kinds of failure factors.

Preliminary Hazard Analysis

The Preliminary hazard analysis (PHA) which is also called rapid risk ranking and hazard identification is used to essentially identify potential accidents related to the system and its interfaces to assess their probability of occurrence and the severity of the damage they may cause and finally propose solutions that will reduce, control or eliminate. Although essential in the process of analysis and safety evaluation of high-risk industrial system, the method PHA is very differently developed and remained unclear. This method is generally classified in theory by an inductive approach [7].

Analysis steps are concluded as follows:

- Hazard identification. Through the experience judgment, the technical diagnosis and so on, try to find the existence danger and the harmful factor in the system;
- Determine the possible accident type. According to the past experiences and lessons, the influence of dangerous and harmful factors on the system is analyzed, and the possible types of accidents are analyzed;
- Develop a pre-hazard analysis table for identified hazards and harmful factors;

- Determine the hazard levels of hazardous and harmful factors, and schedule them according to hazard levels so as to handle them accordingly as planned;
- Formulate safety countermeasures to prevent accidents.

PHA is the precursor to further risk analysis, and is a macroscopic and qualitative analysis method. In the early stage of project development, PHA has the following advantages: it is simple, economical and effective, can provide guidelines for project development, group analysis and design, and identify potential risks, and implement improvements with very little cost and time. The method is generally used in the early stages of the project. When an approximate hazard and potential accident analysis is desired, PHA can also be used to analyze the current equipment.

3.2 Quantitative Safety Assessment Methods

Methods of quantitative safety assessment is the use of a large number of experimental results and extensive accident data statistical analysis of the obtained index or rule based on the (mathematical model), quantitative calculation on the production system of technology, equipment, facilities, environment, personnel management and other aspects of the situation. Safety evaluation results are quantitative indicators, such as the probability of accident, the scope of damage (or hazard), quantitative risk, accident related degree or importance degree of accident caused.

Event Tree Analysis (ETA)

Event tree analysis (ETA) is a forward, bottom up, logical modeling technique for both success and failure that explores responses through a single initiating event and lays a path for assessing probabilities of the outcomes and overall system analysis [8].

This analysis technique is used to analyze the effects of functioning or failed systems given that an event has occurred [9]. ETA is a powerful tool that will identify all consequences of a system that have a probability of occurring after an initiating event that can be applied to a wide range of systems including: nuclear power plants, spacecraft, and chemical plants. This Technique may be applied to a system early in the design process to identify potential issues that may arise rather than correcting the issues after they occur [10]. However, the method cannot analyze the consequences of parallelism, cannot be analyzed in detail, and does not allow conditional independence to be discussed on the event tree, and the size of the event tree grows exponentially with the number of variables. When the quantitative data is limited, a lot of work needs to be done, and ETA is prone to omissions and errors when used in large systems.

Fault Tree Analysis (FTA)

Fault tree analysis (FTA) is a safety technique from engineering developed for analyzing and assessing system safety by uncovering safety flaws and weaknesses of the system [11]. Fault tree analysis (FTA [12]) is commonly used to analyze system safety. It breaks down system level hazards to failures of components, called (failure) events. The events are connected through gates, indicating if all or only any sub-event is necessary to cause the failure. Each event is analyzed recursively, resulting in a tree of events. The branches of a fault tree describe failures of basic components which in combination cause the system hazard.

A fault tree diagram is a logical causality diagram that displays the state of the system (the top event) according to the status of the components (basic events). Like the reliability diagram (RBDs), the fault tree diagram is also a graphical design approach and is an alternative to the reliability diagram [8]. The basic symbols of the fault tree are as follows (Tables 2 and 3):

Table 2. Event symbols.















Name	Symbol	Interpretation
Basic event		The circle symbol represents the basic event which does not need to be verified
Undeveloped event		The diamond symbol represents an event about which insufficient information is available, or which is of no consequence. It also represents omitting events, which generally represent events that may occur but have very small probability values, so it can be ignored in the analysis and calculation
Top event Intermediate event Result event	  	The rectangle symbol represents a top event, an intermediate event, or a result event, that is, an event that needs to be analyzed. Top event is an event of concern in fault analysis, so it must be clear and unambiguous. Intermediate event is an event between top event and result event
External event		The house shape symbol represents external events, the event is bound to or will not occur under normal working conditions. When the conditions in the house shape are satisfied, other inputs of the room shape door are retained, otherwise removed. According to the requirements of the fault can be a normal event, can also be a fault event
Conditioning event		The elliptical symbol represents the conditioning event, which is a condition that restrict or affect logic gates

Table 3. Logic symbols.

Name	Symbol	Interpretation
AND gate		AND gate indicates that the output event occurs when all the input events have occurred.
OR gate		OR gate indicates that the output event occurs when at least one input events have occurred.
NOT gate		NOT gate indicates that the output event is the opposite event of the input event.
Voting gate		Voting gate indicates that the output event occurs only if there are k or more than k events of the N input event have occurred.
Inhibit gate		The elliptical symbol represents the inhibit gate. The output event occurs if the input events have occurred under an enabling condition specified by a conditioning event.
Transfer symbols		Transfer symbols are used to connect the inputs and outputs of related fault trees, such as the fault tree of a subsystem to its system.
Transfer in		
Transfer out		

FTA provides a comprehensive, concise, and visual description of the various factors and logical relationships that lead to disaster accidents, its facilitates the identification of inherent or potential risk factors within the system and provides a scientific basis for design, construction and management; FTA is easy to perform logical calculation, qualitative and quantitative analysis and systematic evaluation. However, there are many steps in FTA calculation; the calculation is more complicated, a lot of work needs to be done when quantitative data are less; large fault trees are difficult to understand and are prone to omissions and errors when used in large systems.

Risk Evaluation Matrix

A Risk matrix is a matrix that is used during risk assessment to define the level of risk by considering the category of probability or likelihood against the category of consequence severity. This is a simple mechanism to increase visibility of risks and assist management decision making [13]. Risk matrix evaluation method is the development of behavior and affected by environmental characteristics or conditions to form a matrix to establish a direct causal relationship between development and environmental impact, method of quantitative or semi quantitative description of the environmental impact of the construction project. Hill first used matrix method in environmental impact assessment in 1966. In 70s, Junior Leopold (Leopold) and others developed the matrix method.

Risk matrix evaluation method is often used in a two-dimensional table to analyze the risk of semi qualitative. Its advantage is that it is simple and quick to operate, so it is widely used. Because this method reveals the level of risk of systems, sub-systems and equipment in sequence, so it difficult to miss any one, and according to the classification and severity of the risk measures are taken in the order of priority, which is more suitable. However, this method is relatively subjective, and if experience is insufficient, it will cause trouble for analysis, and the risk level and frequency of occurrence are determined by the researchers themselves, which may cause a great subjective error.

Analytic Hierarchy Process

The analytic hierarchy process (AHP) is a structured technique for organizing and analyzing complex decisions, based on mathematics and psychology. It was developed by Saaty in the 1970s and has been extensively studied and refined since then. It has particular application in group decision making [14], and is used around the world in a wide variety of decision situations, in fields such as government, business, industry, healthcare, shipbuilding [15] and education.

There are many kinds of risk factors in the process of urban rail transit operation, coupled with the lack of experience in risk identification and related statistical data, the use of analytic hierarchy process can be a good way to avoid deficiencies in this regard. But the result of this method is one kind rough scheme sorting, so for decision making problems with higher quantitative requirements, the application of AHP alone is not appropriate. And it is of much subjectivity, mistakes in judgment can lead to mistakes in decision making.

Monte Carlo Method

Monte Carlo methods (or Monte Carlo experiments) are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results. Their essential idea is using randomness to solve problems that might be deterministic in principle. They are often used in physical and mathematical problems and are most useful when it is difficult or impossible to use other approaches. Monte Carlo methods are mainly used in three distinct problem classes [16]: optimization, numerical integration, and generating draws from a probability distribution.

Monte Carlo methods are widely used in engineering for sensitivity analysis and quantitative probabilistic analysis in process design. The need arises from the interactive, co-linear and non-linear behavior of typical process simulations [16]. However, the simulation system used in practice is very complex, it is difficult to establish the model which must be the collective wisdom, and many uncertain factors must be given in the form of quantitative probability distribution, which makes the real-time operation is difficult. Without calculating the interaction between risk factors, the risk estimation result may be smaller than the reality.

Fuzzy Comprehensive Evaluation Method

In mathematics, fuzzy sets are sets whose elements have degrees of membership. Fuzzy sets were introduced by Lotfi A. Zadeh [17] and Dieter Klaua [18] in 1965 as an extension of the classical notion of set. At the same time, Sali (1965) defined a more general kind of structure called an L-relation, which he studied in an abstract algebraic context. Fuzzy relations, which are used now in different areas, such as linguistics (De Cock, Bodenhofer & Kerre 2000) decision-making (Kuzmin 1982) and clustering (Bezdek 1978), are special cases of L-relations when L is the unit interval $[0, 1]$.

This method gives a mathematical model, which is simple and easy to master. It is a good method to judge the complex problems of many factors and levels, and has a wide applicability. However, because the determination of membership function or membership degree, the evaluation factors and the weight of the evaluation objects are very subjective, so the results are subjective. Moreover, the calculation is more complicated for multi factors and multi-level evaluation.

Neural Network Method

Artificial neural networks (ANNs) or connectionist systems are computing systems inspired by the biological neural networks that constitute animal brains. Such systems learn (progressively improve performance) to do tasks by considering examples, generally without task-specific programming.

This method has a strong learning ability, fault resistance and parallelism. However, under the condition that the known data is insufficient or the training sample set cannot be constructed accurately, it needs to be combined with other comprehensive evaluation methods. In addition, this method is very large in computation. When the sample size is large and the number of neurons in the neural network is bigger, this method is also faced with the danger of finding the global extremum.

3.3 Comprehensive Safety Assessment Methods

Each method has its advantages and disadvantages. When the systems and problems are complicated, one method may not meet the requirements well. Comprehensive safety assessment methods are available.

Fuzzy Analytic Hierarchy Process

Fuzzy analytic hierarchy process method (FAHP) is to classify the evaluation factors according to certain rules to form a hierarchical structure, which usually has three layer structure (general target layer, rule layer and factor layer), and the fuzzy comprehensive evaluation method only a direct target layer and a factor layer. FAHP evaluation method is a combination method of fuzzy comprehensive evaluation (FCE) and Analytic Hierarchy Process (AHP). This method has a wide application in the evaluation system, performance evaluation, system optimization, which is a combination of qualitative and quantitative model, and it is generally the first to determine factors with AHP, then using fuzzy comprehensive evaluation to determine the evaluation results. The fuzzy method is based on the hierarchy method, and the two are mutually integrated and have good reliability to the evaluation.

So, this method has the advantages of FCE and AHP, and also overcomes the weakness of subjectivity in fuzzy analysis. But it also has the disadvantages of both the two methods.

Fuzzy Fault Tree Analysis

In conventional fault-tree analysis, the failure probabilities of system components are treated as exact values. For many systems, however, it is often difficult to evaluate the failure probabilities of components from past occurrences, because the environments of the systems change. Moreover, we often need to consider failure of components which have never failed before [19]. The groundbreaking study of fuzzy fault tree analysis begins with Tanaka and other scholars, take into account the uncertainty of the bottom events and the occurrence probability of top events in the fault tree analysis. The probability of accurate fuzzy probability instead of the traditional reliability analysis, and according to the extension principle in fuzzy mathematics, the components of each system failure (bottom events) of the fuzzy probability to trapezoidal fuzzy number, fuzzy number (multiplication) by approximate calculation, and then based on the traditional definition of importance a function to depict each bottom event to the top event of the contribution, so as to determine the fuzzy important degree of bottom events [20].

This method which has the advantages of fuzzy comprehensive evaluation and fault tree is, it can avoid the strong dependence on the statistical data, and provides a new way to estimate the probability of the accident. Also this method has the disadvantages of the two.

4 Conclusion

In this paper, we review and compare the advantages and disadvantages of various security assessment methods from qualitative, quantitative and comprehensive aspects, which are practical in real work. And these methods give more choices for safety assessment on urban rail transit.

No matter what kind of security assessment method for urban rail transit, it is more important to meet the actual needs of the project. Combinatorial analysis may be better for solving problems, but at the same time the model is more complex and even difficult to solve. In summary, the choice of safety assessment methods must be combined with the actual situation.

Acknowledgments. The project is supported by Tongji University open foundation for traffic science and Engineering (Grant No. 2016J012304), and by Shanghai Metro Line 17.

References

1. Elms, D.: Rail safety. *Reliab. Eng. Syst. Saf.* **74**(3), 291–297 (2001)
2. Marinos, L.: Risk management and risk assessment at ENISA: issues and challenges. In: *International Conference on Availability, Reliability and Security*. IEEE Computer Society, pp. 2–3 (2006)
3. Mahajan, V.: The Delphi method: techniques and applications. *J. Mark. Res.* **13**(3), 317–318 (1976)
4. Kading, L.: *Study and Application of Safety Guarantee System for Urban Rail Transit System*. China Architecture & Building Press (2011)
5. Rydenfält, C., Ek, Å., Larsson, P.A.: Safety checklist compliance and a false sense of safety: new directions for research. *BMJ Qual. Saf.* **23**(3), 183 (2014)
6. Hunt, J.E., Price, C.J., Lee, M.H.: Automating the FMEA process. *Intell. Syst. Eng.* **2**(2), 119–132 (1993)
7. Hadj-Mabrouk, H.: Preliminary hazard analysis (PHA): new hybrid approach to railway risk analysis. *Int. Referred J. Eng. Sci.* **2**(6), 51–58 (2017)
8. Clemens, P.L., Simmons, R.J.: System safety and risk management. NIOSH instructional module, a guide for engineering educators, pp. IX–3–IX–7. National Institute for Occupational Safety and Health, Cincinnati (1998)
9. Wang, J.X., Roush, M.L.: *What Every Engineer Should Know About: Risk Engineering and Management*, p. 264. Marcel Dekker Inc., New York (2000)
10. Ericson, C.A.: *Hazard Analysis Techniques for System Safety*. Wiley, Hoboken (2005)
11. Thums, A., Schellhorn, G.: Model checking FTA. In: *FME 2003: Formal Methods, International Symposium of Formal Methods Europe, and Proceedings, Pisa, Italy, 8–14 September 2003*, pp. 739–757. DBLP (2003)
12. Haasl, D.F., Roberts, N.H., Vesely, W.E. et al.: *Fault tree handbook*. Technical report NUREG-0492, U.S. Nuclear Regulatory Commission (1981)
13. Anthony Jr., L.: What’s wrong with risk matrices? *Risk Anal.* **28**(2), 497–512 (2008)
14. Saaty, T.L., Peniwati, K.: *Group decision making: drawing out and reconciling differences*. RWS Publications, Pittsburgh, p. 8 (2008)
15. Saracoglu, B.O.: Selecting industrial investment locations in master plans of countries. *Eur. J. Industr. Eng.* **7**(4), 416–441 (2013)
16. Kroese, D.P., Brereton, T., Taimre, T., Botev, Z.I.: Why the Monte Carlo method is so important today. *WIREs Comput. Stat.* **6**, 386–392 (2014)
17. Alkhazaleh, S., Salleh, A.R.: Fuzzy soft multiset theory. In: *Abstract and Applied Analysis*, vol. 20 (2012)

18. Alkhazaleh, S., Salleh, A.R., Hassan, N.: Soft multisets theory. *Appl. Math. Sci.* **72**(5), 3561–3573 (2011)
19. Tanaka, H., Fan, L.T., Lai, F.S., et al.: Fault-tree analysis by fuzzy probability. *IEEE Trans. Reliab. R-* **32**(5), 453–457 (1983)
20. Bu, Q., Bi, J., Yuan, Z., et al.: Improvement of Fault Tree Analysis with Fuzzy Mathematical Method (2008)



Research on Application of PC Technology in Shanghai Metro Line 17

Yizeng Wang³, Xiaoqing Zeng², Nixuan Ye², Yining Chen²,
and Zhongzheng Ma¹(✉)

¹ Shanghai Metro Line 17 Development Corporation, Kunshan, China
mazhongzheng2010@163.com

² College of Transportation Engineering, Tongji University, Shanghai, China

³ Shanghai University, Shanghai, China

Abstract. With the continuous development of Chinese modern society, people are no longer only satisfied with the basic material needs, but the pursuit of the spiritual world is also growing stronger. While providing convenient travel services for the citizens, the Shanghai Metro is also committed to letting people enjoy the artistic creativity and ingenuity in the process of operating. Metro Line 17 is such a line full of artistic sense and historical landscape. The works of art, ceramics installations, large space and large vision of Zhuguang Road Station, two murals and PC technology of North Xujing Station were recorded and the observed information was summarized and considered through the investigation activities by taking the Shanghai Metro Line 17 on the spot in this paper. PC is not only the innovation of structural technology, but also the innovation of architecture and art.

Keywords: Metro Line 17 · PC · Culture · Art

1 Introduction

Metro Line 17 is an important Metro line planned and constructed during the 12th Five-Year Plan period in Shanghai, which is from Hongqiao Railway Station in the East (This station is the cross-platform station with Line 2), goes westward parallel along the south side of Songze Avenue, crosses G15 Shenhai Expressway, then transfers along Yinggang East Road and Yinggang Road, goes westward into Qingpu District and Dianshan Lake New Town, and enters Zhujiajiao area, then goes parallel to Oriental Green Boat along the south side of 318 National Road. The construction of the main line was started in 2013 and 2014, with a total construction period of 48 months.

The whole line was operated by Shanghai Metro Second Operating Co., Ltd. at the end of December 2017. The total length of the line is 35.3 km from Hongqiao Transport Hub to Oriental Green Boat. The whole line is set up with Hongqiao Railway Station (transfer with Lines 2 and 10), Zhuguang Road Station, Panlong Road Station, Xuying Road Station, Xujing North City Station, Jiasong Suffer Road Station, Zhaoxiang Station, Huijin Road Station, Qingpu New Town Station, Caoying Road Station, Dianshan Lake Road Station, Zhujiajiao Station and Oriental Green Boat Station. There are 13 stations in total along the line. There are some important areas are

connected, including Zhujiajiao Station, Qingpu New Town, Zhaoxiang Commercial District, Xujing Town, National Convention and Exhibition Center, Hongqiao Transport Hub and other important areas.

2 Create a Public Art of Cultural Landscapes

The Station decoration design of Shanghai Metro Line 17 takes “Lingxiu water village, the source of Shanghai” as the concept theme from the modern water village culture as the breakthrough point, making Line 17 the most beautiful scenery line in Shanghai, reflecting the cultural characteristics of Qingpu and demonstrating the new image of Shanghai as an international city. There are 16 distinct works of art in eleven stations, and each one reflects the local regional characteristics.

2.1 Zhuguang Road Station

Zhuguang Road Station is the second station of Shanghai Metro Line 17 and is located in the urban planning green belt on the south side of Songze Avenue in Qingpu District. The east side of the station is close to the National Convention and Exhibition Center. The opening date is December 30, 2017.

2.1.1 Works of Art

In the white wall, each piece of Shenyao artist’s porcelain artworks and glaze-changing works of Shenyao show elegant artistic style and also become a window of art and culture for people to enter Qingpu District and understand Qingpu District (Figs. 1 and 2).

2.1.2 Zhuguang Kaiwu

“Zhuguang Kaiwu” refers to “knowing the truth of all things under the light of all kinds of wisdom”. Zhuguang Kaiwu is a large-scale pottery installation, which can be regarded as the focus of Zhuguang Road Station. It combines traditional cultural elements such as kaleidoscope, National Convention and Exhibition Center architectural form with four-leaf clover, with frontier scientific and technological elements formed by the abstract patterns such as two-dimensional code, intelligent life, big data, Beidou navigation and so on, and shows the light of nature, the light of humanity, the light of science and innovation, and the light of wisdom. In this way, it implies that the Chinese style is clear, the lights are bright, the national exhibition has four leaves and eight sides are gathered, the tide of science and technology surges, all things are flourishing, the wind and the clouds connect each other, and everything is a new trend (Fig. 3).

2.1.3 Large Space and Large Vision

Usually, the space of the underground railway station is relatively depressed, and the main lighting depends on artificial lighting. Considering the advantages and particularities of the location of Zhuguang Road Station, and considering the improvement of the space effect of existing underground stations, the natural light is reasonably



Fig. 1. Pomegranate pattern ceramic artwork



Fig. 2. Zhuguang road station

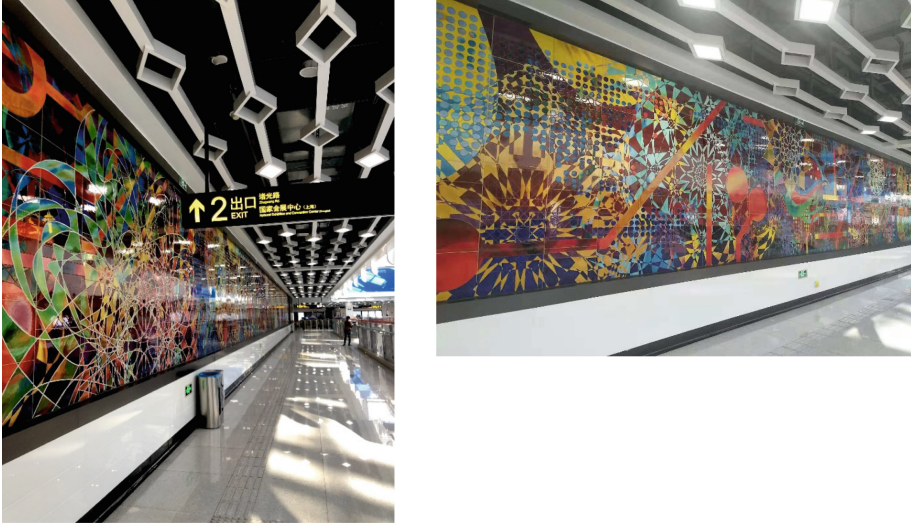


Fig. 3. Large ceramic device “Zhuguang Kaiwu”

introduced into the atrium of the underground station, which is introduced into the station through the lighting skylights and the setting of sunken square on the south side of the station, which optimizes the building space and embodies the concept of energy saving, environmental protection and people-oriented. It can create a natural and harmonious space effect, which provides a reference for the promotion and application of underground atrium station of the urban Metro in the future.

The natural light of the sky is introduced in the transparent atrium design, and the empty glass roof design of tens of meters makes it has a strong spatial sense. When people look up, the clear sky is as blue as being washed, the night sky is as dark as the ink (Fig. 4).

2.2 Xujing North City Station

2.2.1 Starry Sky Relief

The vast starry sky of the universe is like a long river of the years, shining the light of human civilization, recording the marks of human development activities. Fuquan Mountain site culture is a dazzling star in the vast starry sky, reflecting the Majiabang, Songze and Liangzhu cultures. This pattern is based on the starry sky. The relief paintings show the scenes of early human production activities in Fuquan Mountain of Qingpu District and echoes with the local original graphics, original symbols and star projections in the background. The whole scene seems like the dream across the space-time (Fig. 5).

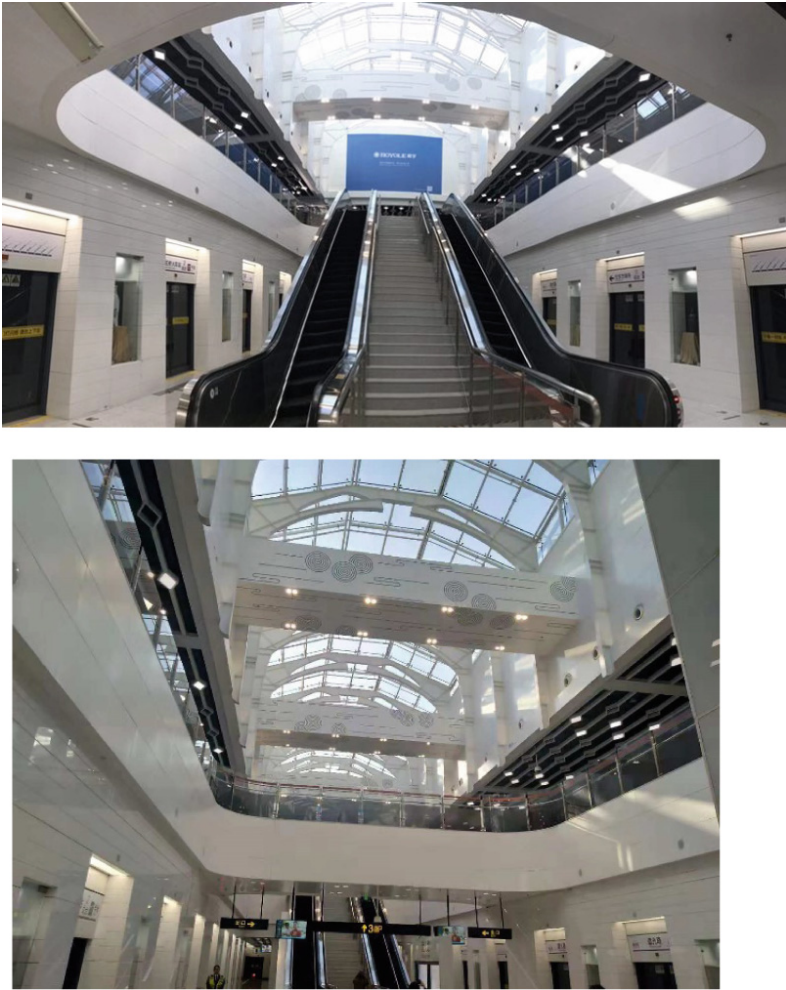


Fig. 4. Large space and large vision

2.2.2 Urban and Rural Mural of Qingpu District

The artwork shows the rhythmic changes of urban and rural rhythm in Qingpu District. From the near-sighted water village to the silhouette of the far-sighted city, through the changeable colors and the scattered layout of the space, it depicts the fresh, elegant and green poetry and painting, and expresses the charm of Qingpu District, which is a town of goodness in the south of the Yangtze River (Fig. 6).



Fig. 5. A group photo at the front of the starry sky relief painting



Fig. 6. Urban and rural mural of Qingpu District

3 PC

3.1 Advantages of PC Assembly Technology

The harmonious appearance of the city requires all buildings to be the corresponding items. At present, our country's Metro stations are still dominated by pouring on site with a very simple shape, and the appearance of the stations is highly homogeneous.

The emergence of PC technology just meets the needs of diversified station models. PC assembly technology has the following advantages:

- (1) Prefabricated concrete components can achieve the same life as the main structure, and there is no need to replace the building components during the valid using period of the main structure.
- (2) Prefabricated concrete components can achieve the maintenance-free. Because of the durability of concrete materials and the related treatment process of the surface coating of prefabricated components and the design of waterproof and water guiding joints, the probability of the corresponding operation and management problems can be greatly reduced, and the cost of operation and maintenance in the later period can be reduced.
- (3) Prefabricated concrete assembly technology can reduce engineering costs. The production cost of prefabricated concrete components mainly lies in the reuse rate of molds. The Metro project has the characteristics of high repeatability and dimension modularization of building components, which can effectively increase the turnover times of prefabricated component formwork and reduce production costs.
- (4) Prefabricated concrete assembly technology can improve the construction quality and reduce the construction cycle. Prefabricated concrete components are prefabricated in a factory at one time. Only a few wet operations or direct assemblies are needed on site, which greatly reduces the construction quality problems caused by manual construction and shortens the corresponding construction period.



Fig. 7. Application of PC technology in Oriental Green Boat Station

- (5) Prefabricated concrete components can achieve a variety of decorative effects and achieve the design requirements of personalized decorative surface of prefabricated concrete components by using different types of formwork.

3.2 Station External Wall Decoration by Using PC Assembly Technology

For the first time, the installation process of prefabricated assembled external wall panels was used in Oriental Green Boat Station, Zhujiajiao Station and Xujing North City Station on Line 17. The large-scale hollow-out board of the platform layer of Oriental Green Boat Station adopts the combination of silica gel and steel film. The inner hole size of each hollow board is different. The solid board and the foot-shank board of the platform layer are selected as texture modeling. After the color treatment, it will become another beautiful scenic line of the scenic spots of Oriental Green Boat in the future (Figs. 7 and 8).



Fig. 8. Application of PC technology in Xujing North City Station

4 Summary

The stations of Shanghai Metro Line 17 combine the regional water village culture. The decoration design takes “Lingxiu water village, the source of Shanghai” as the concept theme to create a cultural landscape with rich water village characteristics. The factory-prefabricated PC standardized materials are applied on a large scale for the first time. A variety of visual effects are created by a simple method of prefabricated PC, just like the Lego building blocks, they show a rich and colorful personalized appearance. Line 17 breaks people’s impression that metro stations are monotonous and have little to highlight. It has become a bright new business card of the city, reflecting the spirit of Shanghai as an international design capital and an innovative city of the times.

References

1. Wang, W., Zeng, X., Shen, T., Liu, L.: Energy-efficient speed profile optimization for urban metro with considerations on train length. In: The 21st IEEE International Conference on Intelligent Transportation Systems. IEEE, Maui, Hawaii, USA (2018)
2. Wang, G., Zeng, X., Bian, D., et al.: Research on modern tram auxiliary safety protection technology based on obstacles detection. In: Proceedings of the International Symposium for Intelligent Transportation and Smart City. Springer (2017). EI:20171603584789
3. Yuan, T., Zeng, X.: Decision analysis of bus travel behavior based on the theory of planned behavior. In: Proceedings of the CICTP 2017 Intelligent Transportation and Smart City 17th COTA International Conference of Transportation Professionals, Reston, VA, USA. ASCE - American Society of Civil Engineers, 7–9 July 2017. EI:17552035
4. Wang, G., Zeng, X., Yuan, T. (eds.): Study on the influence of train control system on service quality of Metro. In: International Conference on Service Systems and Service Management (2017)



Research on Test of Shear Strength of Aggregates Based on Gradation and Particle Shape

Chang-xuan He^(✉) and Yan-feng Bai^(✉)

Shanghai Municipal Planning and Design Research Institute Co., Ltd.,
Shanghai 200030, China
42670900@qq.com, 108359286@qq.com

Abstract. In this paper, aggregate suspension (AC-20I), SUP-19, stone-to-stone contact and interstitial structure (AM-20) and stone-to-stone contact and close-grained structure (SMA-20) were used for samples; contrast research about shear strength of asphalt mixture for four typical types was done through test. Large-size direct shear apparatus for coarse aggregates was used; shear strength affected by aggregates of sphere-shaped and ellipsoid-shaped was discussed. As a result, for aggregate gradation, SMA-20 yielded significantly higher shear strength when compared with the other types; there was no great difference on the shear strength of AM-20 and SUP-19; but AC-20I yielded lowest shear strength. For aggregate shape, aggregate shape was most sensitive for SMA-20 and SUP-19. Shear strength for stone-to-stone contact and close-grained structure was undertaken by coarse aggregate. Mixtures filled by fine aggregate redundantly could develop lower shear strength than unfilled.

Keywords: Gradation · Particle shape · Shear strength of stone-to-stone contact structure

1 Introduction

The shear strength for asphalt mixture can be structured by two aspects: (a) cohesion produced by asphalt binder and (b) internal friction produced by aggregate, which interlock and friction each other. The internal friction is mainly decided by aggregate gradation, particle shape and superficial makings of particle. As discuss, Wang (2005) evaluate the friction properties of aggregates using analogue principle. The virtual testing based discrete element method is also introduced to evaluate the shearing characteristic (Wang 2003). In this paper, large-size direct shear apparatus is used for research; particle shape and aggregate gradation are considered.

2 Experimental Programme

The laboratory work consisted of testing various granular materials which included: (a) SMA-20 (b) AM-20 (c) SUP-19 (d) AC-20I (e) coarse aggregate composing SMA-20. The materials were tested under various particle shapes: sphere-shaped and ellipsoid-shaped.

2.1 Material Preparations

Particle shape can be described by three-dimensional size of one particle. Coarse aggregates (>4.75 mm) were obtained from Wenjiang. Bank gravels were considered to be suitable for testing. Sphere-shaped ($a \approx b \approx c$) and ellipsoid-shaped ($a/c \leq 2.5$ and $b/c \leq 1.5$) were researched in this paper. The particle size distribution is shown in Table 1.

Table 1. Particle size distribution of four typed mixtures

Sieves size (mm)		The percent passing the sieves (%)											
		26.5	19	16	13.2	9.5	4.75	2.36	1.18	0.6	0.3	0.15	0.075
SMA-20	Required by code	100	90	72	62	40	18	13	12	10	9	8	8
		100	100	92	82	55	30	22	20	16	14	13	12
	Used	100	95	82	72	47.5	23.4	18.9	14.9	12.6	11.9	11.2	9.5
AM-20	Required by code	100	90	60	50	40	15	5	2	1	0	0	0
		100	100	85	75	65	40	22	16	12	10	8	5
	Used	100	90	60	50	40	15	10	6.7	4.1	2.6	2.1	1.5
SUP-19	Used	100	95.0	85.0	74.5	63.5	42.5	28.7	17.1	9.9	6.8	5.7	4.5
AC-20I	Required by code	100	95	75	62	52	38	28	20	15	10	6	4
		100	100	90	80	72	58	46	34	27	20	14	8
	Used	100	97.5	87.5	77.5	68.5	57.8	41.6	29.0	17.6	10.3	7.5	5.9

2.2 Testing

A series of direct shear tests on researched gradation and particle shape was carried out in direct shear box having dimensions of 50 × 40 cm (d × h). The large box is suitable for aggregates containing particle up to 100 mm. The tests used four different vertical pressure: 40, 80, 120, 160 kN. The material were compacted in the box in three layers. Each layer was compacted using a vibration excitor. In each test the material used fort preparing the sample was dried to constant weight under 105 °C. Sample were sheared at 4 mm/min. During shearing the shear displacement, horizontal displacement and the shear load applied to the sample were carefully recorded. Using this information the required stress-strain parameters were calculated.

2.3 Test Result and Evaluation

2.3.1 Effect of Particle Shape to Shear Strength for Asphalt Mixture

Figures 1, 2, 3 and 4 illustrate the stress-strain relationship for ellipsoid-shaped and sphere-shaped under four different gradations. The samples were taken to the maximum shear displacement, which could be applied in one direction of the direct shear apparatus. This corresponds to a shear strain of approximately 10%. Results indicated that the shear stress increased with increasing shear displacement and reached a peak at this shear strain. The ellipsoid-shaped aggregates exhibited higher shearing resistance than the sphere-shaped specimen. The improvement in the shearing resistance of the ellipsoid-shaped aggregates was mainly due to the interlocking action between one ellipsoid-shaped particle and adjacent ellipsoid-shaped particle. For shearing resistance, aggregate shape was most sensitive for SMA-20 and SUP-19 while least sensitive for AC-20I. Sliding friction on particle surface, interlocking action between one particle and adjacent particle and rearrange of particles after shear failure form shearing resistance. Coarse aggregates form skeleton in SMA mixture, fine aggregates completely fill void among skeleton, consequently, strength needed when rearrange of particles after shear failure is hugely influenced by particle shape. The content for fine aggregates making up of AC-20I mixture (aggregate suspension structure) can reach 60% of mixture weight; coarse aggregates can not touch each other, the influence for particle shape may be ignored.

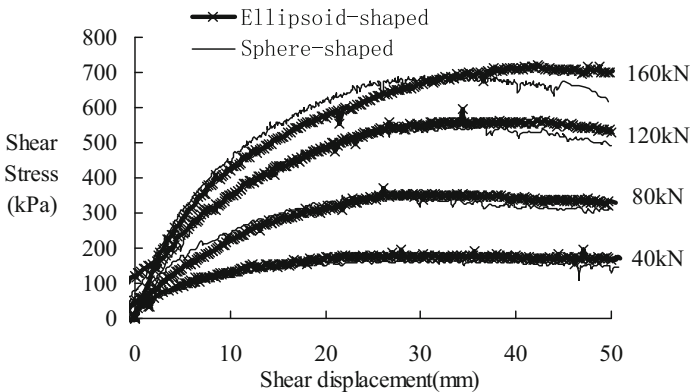


Fig. 1. Stress displacement behavior of AC-20I in large direct shear box

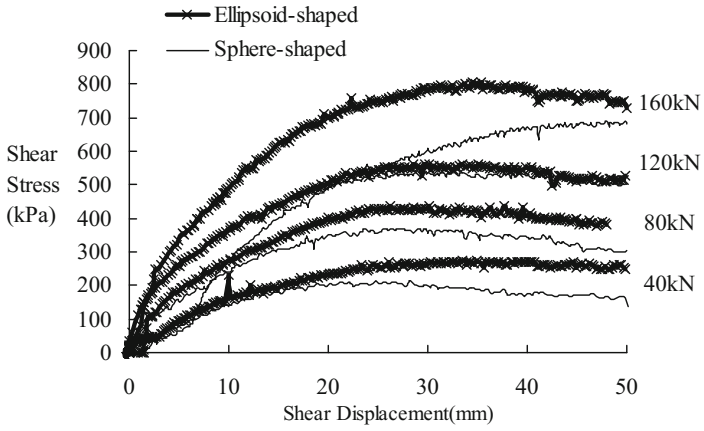


Fig. 2. Stress displacement behavior of SUP-19 in large direct shear box

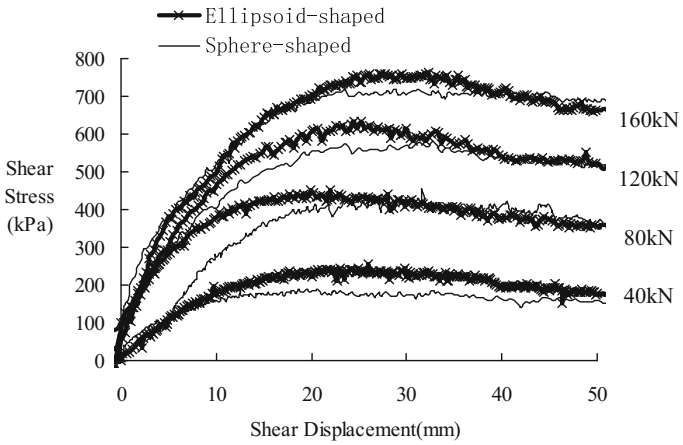


Fig. 3. Stress displacement behavior of AM-20 in large direct shear box

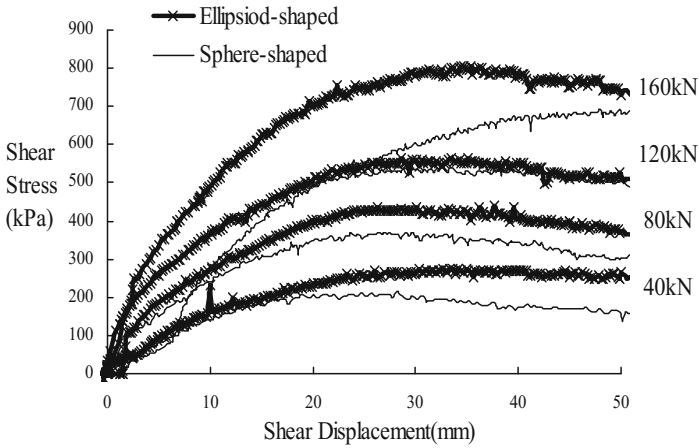


Fig. 4. Stress displacement behavior of SMA-20 in large direct shear box

2.3.2 Effect of Aggregate Gradation to Shear Strength for Asphalt Mixture

Figures 5, 6 and Table 1 show the comparison for the angle internal friction of four typical aggregate gradations under two different particle shapes. The results indicated that for aggregate gradation, SMA-20 yielded significantly higher shear strength when compared with the other types; there was no great difference on the shear strength of AM-20 and SUP-19; but AC-20I yielded lowest shear strength.

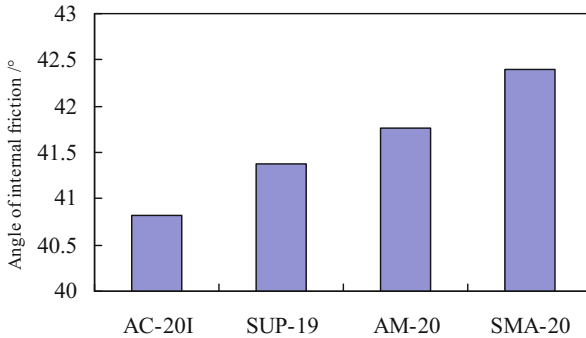


Fig. 5. Comparison for shearing resistance in terms of friction angle (particle shape is sphere-shaped)

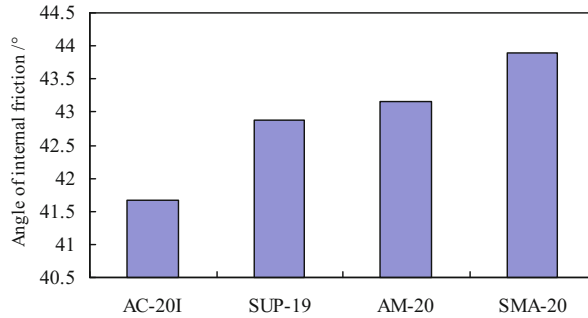


Fig. 6. Comparison for shearing resistance in terms of friction angle (particle shape is ellipsoid-shaped)

Consideration must be given to the Table 2. The friction angle for AC-20 (Ellipsoid-shaped) was 41.66° while AM-20 (Sphere-shaped) was 41.77° , the difference of friction angle of two gradations is small. It also appears that the friction angle for AM-20 (Ellipsoid-shaped) was 43.17° , which reduced to 42.20° for SMA-20 (Sphere-shaped). This phenomenon demonstrated that the shearing resistance for mixtures is determined by the particle sizes and the grading of particle.

Table 2. Comparison for shearing resistance in terms of friction angle

Gradation	Sphere-shaped				Ellipsoid-shaped			
	SMA	AM	SUP	AC	SMA	AM	SUP	AC
Angle of internal friction	42.40	41.77	41.38	40.82	43.89	43.17	42.88	41.66

2.3.3 Discuss for Forming Mechanism of Shear Strength for Stone-to-Stone Contact and Close-Grained Structure

SMA mixture can be considered the most typical sample of stone-to-stone contact and close-grained structure. In respect that SMA-20 was used in this research, the particles whose size is more than 4.75 mm could compose the skeleton of mixtures. Coarse aggregates composing the skeleton were tested under the same condition above.

Figure 7 illustrates the relationship between shear displacement and shear stress for SMA mixture and the skeleton part of SMA mixture. Figure 8 shows the relationship between the maximum shear stress and the corresponding normal stress for SMA mixture and the skeleton part of SMA mixture. They appear that the two slopes of the failure envelopes are not equal. The internal friction of the skeleton part of SMA mixture is greater than which of SMA mixture. This phenomenon can possibly be explained that shear strength for stone-to-stone contact and close-grained structure was undertaken by coarse aggregate, mixture filled by fine aggregate redundantly could develop lower shear strength than unfilled. The shear strength comes from the intercalation between aggregates, fine aggregates cannot provide embedding.

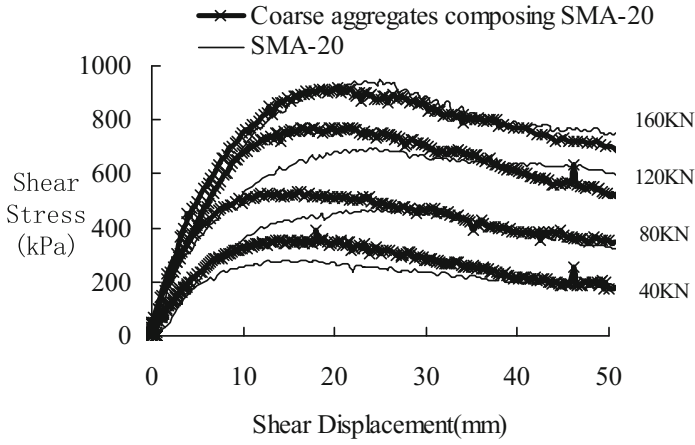


Fig. 7. Stress displacement behavior of skeleton structure in large direct shear box

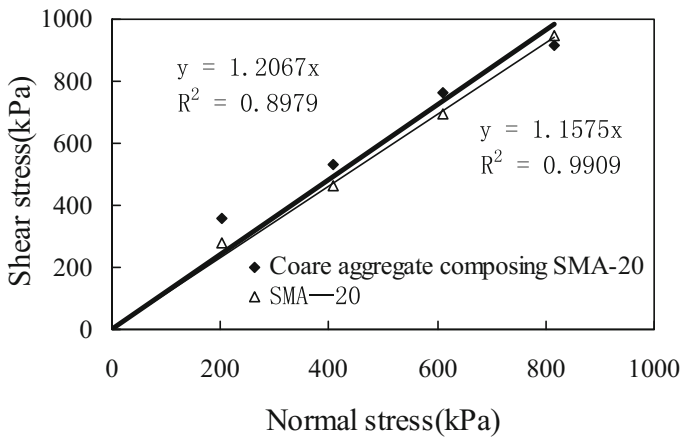


Fig. 8. Failure envelope for skeleton structure

3 Conclusion

Three aspects have been examined in this research: (a) influence of particle shape on the strength characteristic of mixture, (b) influence of aggregate gradation on the strength characteristic of mixture and (c) forming mechanism of shear strength for stone-to-stone contact and close-grained structure. The objective of the study was to examine the shear strength characteristics of various mixtures.

- (1) For particle shape, the ellipsoid-shaped aggregates exhibited higher shearing resistance than the sphere-shaped specimen.
- (2) For aggregate gradation, SMA-20 yielded significantly higher shear strength when compared with the other types; there was no great difference on the shear strength of AM-20 and SUP-19; but AC-20I yielded lowest shear strength.
- (3) The shearing resistance for mixtures is influenced differently by the particle sizes and the grading of particle.
- (4) AM mixture belongs to stone-to-stone contact structure, which is similar to SMA mixture. But there is obvious difference in the interlocking action for SMA mixture and AM mixture.
- (5) Shear strength for SUP-19 mixture is higher since its gradation is uniform.
- (6) Shear strength for stone-to-stone contact and close-grained was undertaken by coarse aggregate. Mixture filled by fine aggregate redundantly could develop lower shear strength than unfilled.
- (7) The performance for asphalt mixtures with different gradations validated the above conclusions.

References

- Wang, D.-y.: Evaluation on frictional behavior of aggregates amongst asphalt mixture with direct shear test. *Central South Highway Engineering*, vol. 30, no. 3 (2005)
- Wang, D.-y.: Evaluation on grading type of asphalt mixture with virtual test method. *J. South China Univ. Technol. (Natural Science Edition)*, vol. 31, no. 2 (2003)
- Technical Specification for construction of highway asphalt pavement (JTG F40-2004). China communication press (2004)
- Specifications for Design of Highway Asphalt Pavement (JTJ041-97). China communication press (1997)
- Yan, J.-j.: *Construction materials for road*. China communication press (1986)
- Liu, C.-y.: *Soil mechanics*, pp. 146–148. China railway press (2002)
- Jiang, L.-W.: Self-organized criticality and the application in granular mixture. Academic Degree Applied for Master Southwest Jiaotong University (2003)



The Seismic Analysis of the Structure of Hydraulic Tunnels

Gang Wang^(✉)

Shanghai Municipal Engineering Design and Research Institute (Group)
Co., Ltd., Shanghai, China
wanggang@smedi.com

Abstract. Recently, large hydraulic tunnels keep appearing since the urbanization develops. According to the earthquake damage records, the seismic design of these large hydraulic tunnels has been gradually taken seriously. Based on several conventional anti-seismic theories, the mechanical behaviors under earthquakes are analyzed. However, there are still many differences in hydraulic tunnels from traditional tunnels, such as their force conditions, operating requirements and joints structures. In the end, several special problems of hydraulic tunnels are proposed in this paper, which may provide reference for further research.

Keywords: Hydraulic tunnel · Seismic design

1 Background

Recently, all kinds of super long and large hydraulic tunnels keep appearing since the urbanization develops. These tunnels are built to solve many city problems, such as waterlogging, contamination prevention and waste water transferring.

Decades ago, underground structures were usually considered with better anti-seismic abilities. However, the earthquake in southern Hyogo prefecture (Japan) in 1995 broke quite a few underground rail stations. It overthrew the concept that underground seismic calculation can be ignored. In China, traditional hydraulic pipes are relatively small and simple, seismic design was usually not taken seriously. But in recent years, the existing water supply and sewage pipes have gradually approached their service life, and even some of them have been used under over loads. Meanwhile, new hydraulic pipes usually need to be constructed as deep and large tunnels with many important structures in the nearby. As a result, people have to pay more attention to the impacts of earthquakes. Now, the seismology itself has also been developed, the anti-seismic codes and standards have provided new requirements for the pipe seismic design.

Several codes can provide reference for hydraulic tunnels: Code for seismic design of buildings, Code for seismic design of special structures, Code of seismic design of outdoor water supply, sewerage, gas and heating engineering, Code for seismic design of urban transit structures, Code for seismic design of subway structures. Therein, the requirements of displacements and strains have been proposed in the Code of seismic design of outdoor water supply, sewerage, gas and heating engineering, but there are no

specific methods for the cross and longitudinal section calculation in large-diameter hydraulic tunnels. In the Code for seismic design of urban transit structures, there has been the response displacement method and the response acceleration method for cross section calculation, the response displacement method for longitudinal section, and the time history analysis method, which have provided a good many references for the large-diameter hydraulic tunnels for their similar structures. However, there is no water inside the urban transit tunnel, so at least the inertia force of the inner water is not considered. Nowadays, there are many types of large hydraulic tunnels, such as black and odorous water treatment pipes, old hydraulic pipe maintenance pipes, and super deep water storage and drainage tunnels. The scale, importance and complexity of these large-diameter hydraulic tunnels have been no less than urban transit tunnels, so the analysis of their seismic design is no longer ignorable.

2 Theories of Pipe Seismic Calculation

Traditional seismic design theories basically originate from building seismic theory. The earthquake effect can be considered as the inertial force response caused by earthquake stimulation. But the buried pipe is a kind of underground structure, the original soil mass is far more than the pipe structure. Besides, the pipe is restrained by soil, and the earthquake energy is significantly lowered down because of the absorption of the surrounding soil, the pipe can hardly produce resonance as the aboveground structures, therefore, there is no enough pipe acceleration because of earthquake. Koizumi et al. concluded that the response by the soil displacement should be mainly analyzed, and meanwhile, if the pipe is full of water, the inertial force of the inner water may have a certain impact on the pipe anti-seismic behavior.

In the engineering design, the main seismic methods are the quasi-static method, the response displacement method, the response acceleration method and so on. This paper based on the characteristics of hydraulic tunnels and the existing codes recommendation, chooses several methods for introduction.

2.1 Quasi-static Method

It is a method of simplifying the earthquake effects as static forces. For aboveground structures, it can make an hypothesis that all the parts of the structure has the same vibration as the earthquake, and then the inertia force can be calculated as the product of ground acceleration and structure mass. For underground structures, it can simplify the outer soil force as static and dynamic soil pressure, which are both calculable, so the inner force, crack and deformation of the structure can be calculated.

2.2 Response Displacement Method in Lateral Earthquakes

In this theory, the deformation of underground structure is restrained by the soil deformation. Part of the soil deformation is transferred to the structure, causing the stress and strain of the structures. In normal conditions, the strain transferring ratio α and β are used to reflect the decrement of transferring of soil to the structure, so the

axial strain and bending strain can be obtained. In the cross section analysis, the restraining effect of soil can be simplified as springs, and secondly, the calculated earthquake displacement is applied to the lateral springs, and then by the calculated maximum acceleration distribution, the horizontal load and dynamic outer water pressure is obtained, and at last, the maximum shear strength is applied on the top of the structure. The sum of horizontal shear strength and inertia strength is applied on the structure bottom to calculate the structure seismic responses (Fig. 1).

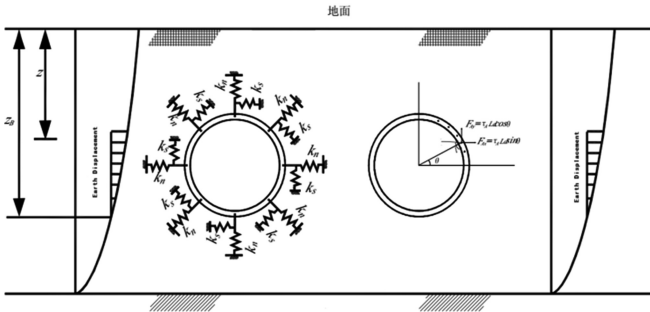


Fig. 1. Diagram of lateral response displacement of tunnels

2.3 Longitudinal Response Displacement Method

When using longitudinal displacement method, the surrounding soil can be treated as the supporting springs, and the structure can be simulated as beams, as shown in Fig. 2. The displacement of soil can be applied to the soil springs. In the seismic design, the maximum lateral displacement can be calculated by the tunnel depth, and through the displacement function along the tunnel axis, and then the tunnel lateral force could be calculated by the lateral constraining stiffness.

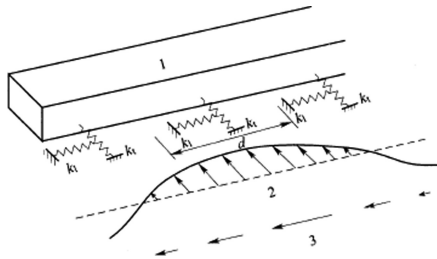


Fig. 2. Diagram of longitudinal response displacement of tunnels

3 Calculation Example

3.1 Model Analysis

Based on a certain shield tunnel, the calculation process by response displacement method is shown as bellows. The inner diameter of the tunnel is 5.5 m, the thickness of the lining is 0.25 m, there is no 2nd lining. The seismic action is set as M8.

The model is calculated by MIDAS GTS. In the model, the soil is simulated by Mohr-Coulomb constitutive model. In the dynamic finite element analysis, the high frequency (shortwave) decides the element length, and the low frequency (long wave) decides the model boundary. Considered the impact of horizontal and vertical earthquake waves, the lateral artificial boundary distance is three times of the structure width, and the bottom artificial boundary distance is set as the earthquake datum, and the upper boundary is set as the earth surface, as shown in Fig. 3.

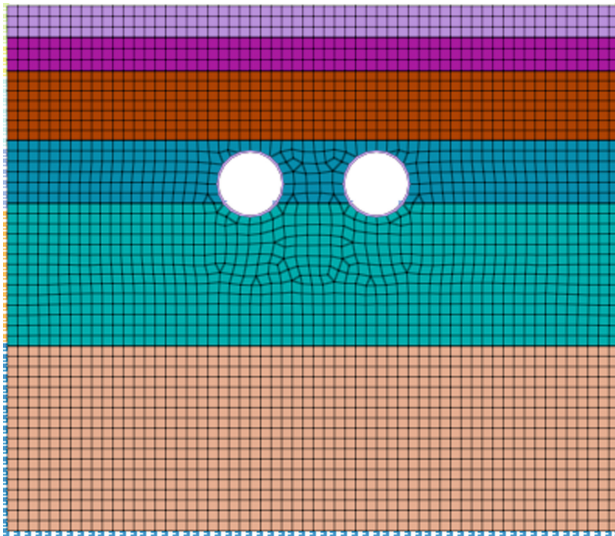


Fig. 3. The calculation model of tunnel and soil layers, two shield tunnels are built in the earth layers, and the boundary of the model is set as more than 3 times of diameters.

3.2 The Analysis of the Seismic Calculation Results

- (1) Lateral displacements under earthquake
 - (a) Structure deformation

Part of the calculation is extracted. Under the design earthquake, the maximum lateral displacement and vertical displacement are shown in Figs. 4 and 5, and the maximum diameter deformation is shown in Fig. 6. The results reveals that the maximum diameter deformation is 5.68 mm.



Fig. 4. The lateral displacement under design earthquake effect

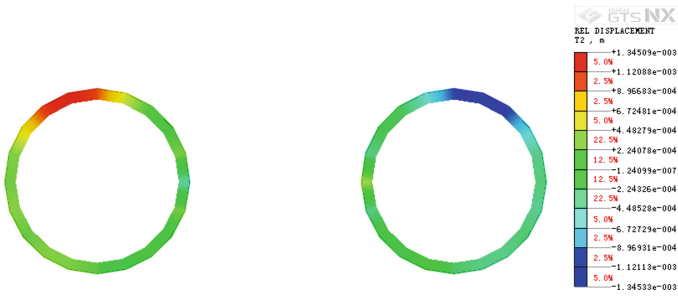


Fig. 5. The vertical displacement under design earthquake effect

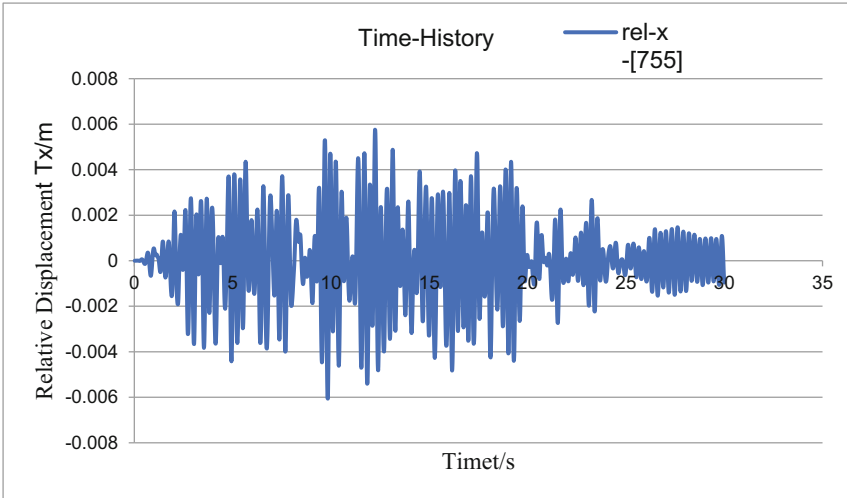


Fig. 6. The diameter deformation-time under design earthquake effect

The maximum diameter of the shield tunnel is concluded as Table 1.

Table 1. The deformation of the tunnel

Section	Diameter deformation (mm)	Maximum diameter deformation ratio
1	5.68	1.03‰

(b) Structure force

The structure force under design earthquake effect is shown in Figs. 7 and 8. The maximum bending moment is 180.68 kN·m, and the corresponding axial force is 810 kN. And the calculation according to these two data show that the structure can meet the requirements in the code.

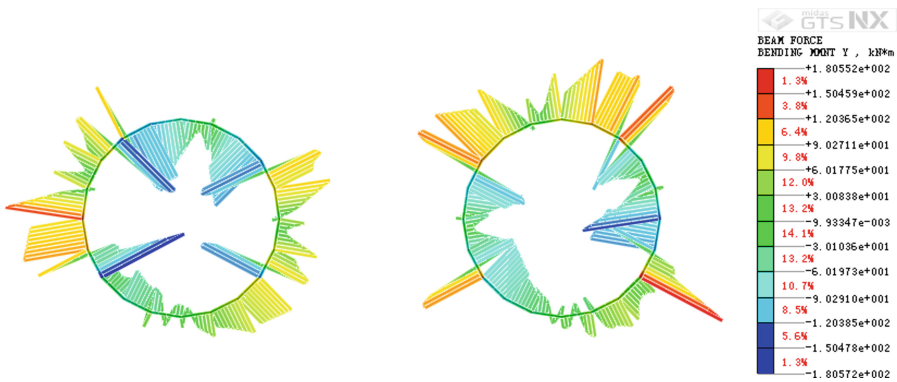


Fig. 7. The bending moment of the structure under design earthquake effect

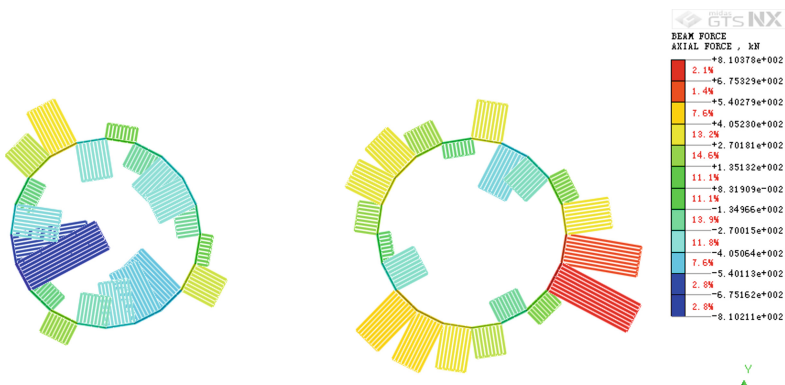


Fig. 8. The axial force of the structure under design earthquake effect

4 Several Problems to Be Solved in Hydraulic Tunnel Seismic Design

The summary and analysis of seismic design of tunnels in China started relatively later than developed countries, the statistic data of large underground tunnel seismic characteristics are also not enough. There are few large disasters caused by earthquakes should thank to the relative strong anti-seismic performance of the tunnels themselves, so many problems are covered. However, more and more large and super large hydraulic tunnels are very different from traditional shield tunnel or jacking pipes without inner water pressure, there are significant differences in the loading principle and the operating requirements, which means the seismic design has to be flawless as possible.

In recent years, seismic design has been taken more and more seriously in traditional large shield tunnels. There are also many seismic design codes and guidelines which are suitable for large pipes. Nevertheless, there are still some special problems to be considered, and these problems happen to be easily to cause serious aftermaths under earthquakes.

(1) Longitudinal axial force

Given the environment and operating requirements, there are usually curve parts in a hydraulic line. Thanks to the development of modern curve shield tunnel technology, some of them already have the ability to make a turning radius of 80D (outer diameter of the tunnel). The problem is that the axial of the tunnel can also generate tremendous force because of earthquake effect, which could cause significant offsets at the curve parts, and consequently, the joints may be invalid, leading to serious leaking.

(2) Inertial force of inner water

The traditional transportation tunnel, utility tunnel and power tunnels are all without inner pressure, the earthquake effects are mainly caused by soil displacements. But there is water inside the hydraulic tunnels, which may generate flow sloshing under earthquake effects, and as a consequence, causing resonance.

(3) The joint of tunnel and shaft

Based on the post-earthquake observation results in Japan, the main structure of tunnels are basically not damaged, but there are many cracks and displacements at the joints of the tunnel and the shafts. The deformation and crack control requirements are much more strict in hydraulic tunnels, so the joint structure need to be specifically considered under seismic design, such as the flexible grouting to the surrounding soil, or using the flexible joints to counteract the deformation or displacements caused by earthquakes.

5 Conclusions and Prospects

The anti-seismic performances of hydraulic tunnels are introduced in this paper. There are quite a few research results and codes for traditional transportation tunnels, and jacking pipes in water supply and sewerage engineering, which may provide references for large hydraulic tunnels. The example in this paper shows how to derive calculation results of hydraulic tunnel by the existing seismic theories.

However, there are many differences between the hydraulic tunnels and other tunnel without inner pressure on the load principle and operating requirements, so the existing design methods may not avoid all earthquake damages. On the basis of observation data and theoretically analysis, this paper proposed several special problems in hydraulic tunnels under earthquakes, which may provide thoughts for further research in the hydraulic tunnel seismic design.

References

1. Koizumi: Seismic research and calculation examples of shield tunnels (2009)
2. Code for seismic design of buildings, GB50011-2010 (2016 version), national standards of People's Republic of China
3. Code for seismic design of special structures, GB50191-2012, national standards of People's Republic of China
4. Code of seismic design of outdoor water supply, sewerage, gas and heating engineering, GB50032-2003, national standards of People's Republic of China
5. Code for seismic design of urban transit structures, GB50909-2014, national standards of People's Republic of China
6. Code for seismic design of subway structures, DG TJ 0820642009, Shanghai standards in construction
7. Japan society of civil engineers, Tunnel standards [Shield tunnels] and illustrations (2011)
8. Zheng, Y.L., Yang, L.D., Li, W.Y., Zhou, J.: Seismic of underground structures. Tongji University Press (2011)



Analysis on the Effect of Excavation by Sections for Large Foundation Pit Without Horizontal Struts

Dongqing Nie^(✉)

Shanghai Municipal Engineering Design Institute (Group) Co., Ltd.,
Shanghai, China

niedongqing@smedi.com, ndqasd@163.com

Abstract. Excavation by sections is usually used in large foundation pits without horizontal struts. The pile inclination monitoring of the excavation of Bailonggang underground wastewater treatment plant shows that the adjacent soil can provide significant restraining effects for the piles of 1st excavation section, and the excavation of the 2nd excavation section has great impacts on the piles in the 1st excavation section. Based on the monitoring result, a numerical calculation model by the finite difference method is built to study the effects of construction sequences and section width. The parametric study showed that the deformation of the excavation can be effectively controlled with the construction of the cushion and raft. The restraining effect to the piles from the cushion is just a little lower than the raft, while the deformation of the pit could be quite large before the cushion is constructed. The deformation of the retaining structure is reduced when the excavation section width is smaller. The analysis in this paper reveals that there is an optimal excavation section width, and the deformation of retaining structure cannot be additionally reduced if the excavation section width is smaller than the optimal width.

Keywords: Large foundation pit · Without horizontal struts · Bailonggang · Excavation by sections · Deformation · Finite difference method

1 Introduction

Large foundation pits are usually excavated by divisions and layers. Therefore, it can not only use the spatial effect of foundation pits to reduce deformation, but also guarantee the stability. The spatial effect of foundation pits has been studied by scholars at home and abroad. Liu et al. proposed time-space effect rule in foundation pits according to the engineering practices in Shanghai. Ou C Y et al., Lee F H et al., Yu et al., Gao et al. researched the 3D spatial effect of foundation pits by FEM methods. Lu et al., Lai et al. analyzed the ring beam influence to the foundation pit spatial effect. Lei et al. deduced the calculation of foundation pit spatial effect coefficient by limit equilibrium method, which based on the top plasticity theory and the associated flow rule of soil. Liu et al., Li et al. studied the influence of foundation pit spatial effect based on foundation pit monitoring. Cheng et al. [11, 12] did research on the load transferring

mechanism of foundation pit with spatial effect when there is partial failure based on the 3D finite-difference excavation model and model test.

There has been research considering the effect of foundation pit deformation and stability, but the foundation spatial effect is rarely considered, especially for those adopted different excavation methods. This paper is on the basis of the foundation pit in Shanghai Bailonggang wastewater treatment plant, and then a three-dimensional numerical model is built, for the analysis of the effect of different construction sequences and excavation width for the foundation without horizontal retaining structures.

2 Research Question Description

Shanghai Bailonggang wastewater treatment plant is located in Heqing, Pudong district of Shanghai. The area of foundation pit is 200000 m², and the excavation depth is 12.8–15.8 m. The elevation of site is 4.3 m (Wusong Elevation System). As the foundation pit is tremendous and deep, and the environment is relatively simple, so the double-row piles without horizontal retaining structures are adopted.

The A-A profile of southwest foundation pit is shown in Fig. 1. The upper foundation is excavated by step-slope method, and the lower part is retained by double-row piles. The front and middle of two piles is strengthened by three axes agitating pile. The two slope heights are both 3.5 m, and the depth of double-row pile is 7.5 m, and the excavation depth is 14.5 m. The front and back piles are both 25 m long, and they are both constructed by Ø1000@1200 bored pile. The front and back piles are connected by coupling beams, which are shown as in Fig. 2. The piles and beams are constructed by Concrete C30. Considered there is a thick layer of silt clay at the depth of excavation face, the soil between and in front of piles is strengthened by three axes agitating piles. The strengthening length between piles is 18 m, and the passive zone of pit is strengthened by boundary area of 5 m at depth and 6 m at width. The waterproof curtain is 5 m outside the slope top, which is constructed by 16 m long three axes agitating piles.

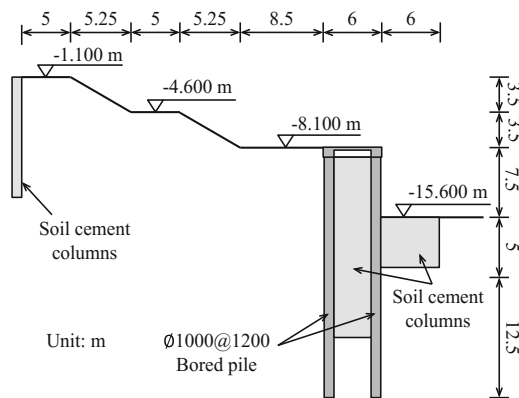


Fig. 1. A-A profile of Bailonggang foundation pit

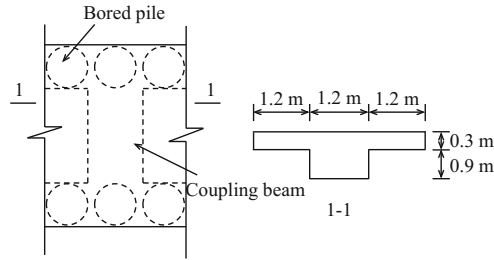


Fig. 2. Coupling beam

The foundation plan is shown in Fig. 3. It is 575 m long and 344 m wide, the excavation depth is 12.8 m in the north and 14.5 m in the south. The pit is excavated by 30 sections, which is shown in Fig. 3. The right part of the pit is symmetrically arranged. Each section is about 30–60 m wide. Inclinator tubes are set at the front piles to monitor the deformation of retaining structure, and the distance of tubes is about 20 m.

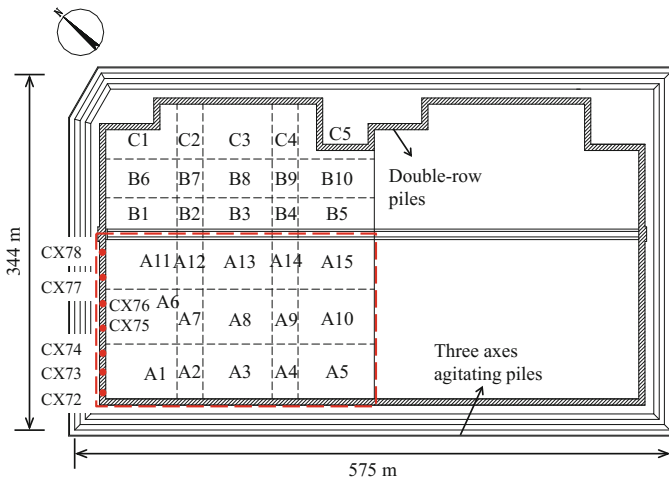


Fig. 3. Foundation plan of the Bailonggang underground wastewater treatment

3 Analysis of Monitoring Results

As the southwest pit excavation depth is 14.5 m, this paper mainly discuss the monitoring data of this area. There are 7 monitoring spots in the southwest, which are shown in Fig. 3. The maximum inclination data are shown in Fig. 4. According to the records, the A1 section was dewatered and excavated on March 5th in 2018, and reached the designed depth on March 15th. The CX72–CX74 showed the inclination

near the corner is smaller because of the spatial effect of foundation pit. The displacement of CX72 at the edge of foundation is less than half of CX76, which is in the middle, so the spatial effect played a great role on the foundation deformation. The A6 section was excavated on April 14th, and reached the designed depth on April 20th, when the raft started to be constructed. As shown in Fig. 4, the maximum displacement of CX72–CX74 was stable until A6 section was excavated. It indicated that the surrounding soil has a certain constraint effect before excavated. To study the impact of excavation by sections, a numerical model of Bailonggang foundation pit is built in this paper, and the effect of construction sequences and section width is analyzed.

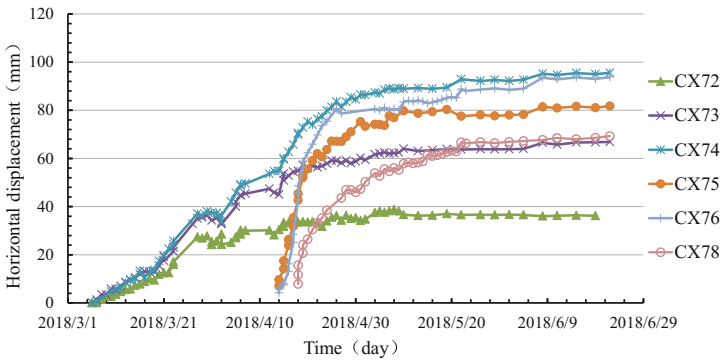


Fig. 4. Maximum of inclination - time

4 Numerical Calculation Model and Parameter Analysis

Based on the Bailonggang excavation pit, a two-dimensional plane strain model is built by finite difference method, which is shown in Fig. 5. The excavation width is set as 120 m to save calculation cost, and the excavation width in the model is set as 60 m. The total width of the model is 150 m, and the height is 60 m. The left, right and bottom of the model are all restrained.

In the finite difference model, Cap-yield model (CY model) is adopted to reflect the hardening characteristics of soil. There are three hardening laws in this constitutive model: the cap-hardening law, the friction hardening law and the compaction/dilation law [13]. This constitute model widely used in the numerical calculation of geotechnical engineering such as tunnels [13, 15] and foundation pits [14]. The parameters of the calculation model are shown in Table 1.

Row piles are simulated as linear elastic piles, and the coupling beams are simulated as lining elements with the bending stiffness of 1 m-thick slabs. The soils between piles and in the pit are simulated by Mohr-Coulomb model. The specific calculation parameters are shown in Table 2.

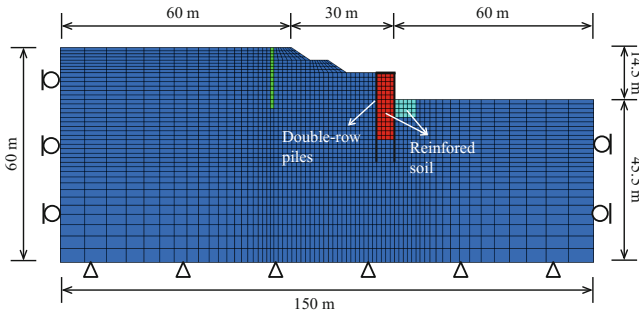


Fig. 5. Finite difference calculation model

Table 1. Mechanical parameters of soil.

Soil layer	Soil name	Depth/m	Dry density/kg·m ³	Compression modulus/MPa	Porosity	C/kPa	$\phi/^\circ$
①3	Dredger fill	1.9	1112	2.1	0.586	11.0	20.9
②31	Grey silty clay	5.4	1400	8.0	0.473	10.0	28.0
②32	Grey sandy silt	9.7	1426	10.3	0.460	6.0	30.5
③	Grey mucky silty clay	13.9	1243	2.9	0.537	8.0	25.9
④	Grey silt clay	24.7	1115	2.2	0.584	10.0	23.2
⑤1	Clay	29.9	1204	2.7	0.552	9.0	24.7
⑤31	Silty clay with clay	33.4	1353	4.3	0.494	7.0	29.1
⑧11	Silty clay and clayey silt	38.7	1370	4.7	0.487	6.0	30.2
⑧12	Silty clay with silt	60.0	1355	4.5	0.492	6.0	29.9

Table 2. Structure calculation parameters.

Name	Element type	Modulus/MPa	Poisson' s ratio
Row pile	Pile	30000	0.2
Coupling beam	Liner	30000	0.2
Strengthening mass	Solid	60	0.3

The comparison of numerical calculation results and site monitoring data is shown in Fig. 6. The displacements of the calculation results and the monitoring data are of the front piles. As the inclination monitoring of the double-row piles is started after two-stage slope is finished, the deformation of slope is ignored in the calculation. In the monitoring, the maximum measuring depth is 18 m below the pile top, so the deformation lower than

18 m is also ignored. The middle spots in the foundation are chosen to make a comparison of calculation results and monitoring data. As shown in Fig. 7, they are basically the same, but the monitoring displacements on the top of the piles are obviously less than the calculation results, which may be caused by the big stiffness of coupling beams, and the foundation pit was excavated by sections, the coupling beams of the sections before excavation may provide constraint for the upper piles.

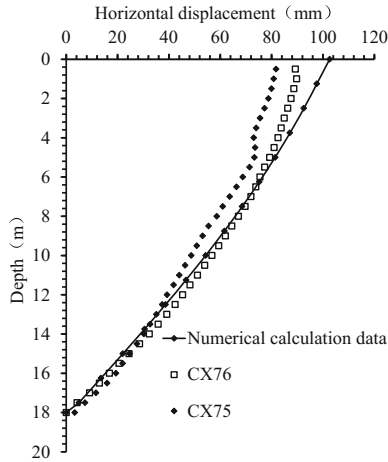


Fig. 6. Comparison of the calculation and the monitoring data

In order to study the effects of excavation by sections, the two-dimensional plan strain model is expanded to three-dimensional calculation model. The foundation pit is excavated by “one and next”, as shown in Fig. 7. A half of the real foundation pit is needed in the model because of symmetry so as to save calculation costs. Half of the first excavation width ($B/2$) and half of the next excavation width ($B/2$) are built in the model. And then the construction sequence and section width are simulated in this model. As the maximum deformation of the retaining structure is usually in the middle of the foundation pit, the analysis mainly focused on the middle piles. The middle pile of first excavation section is named as Z1, and the middle pile of next excavation section is named as Z2, as shown in Fig. 7.

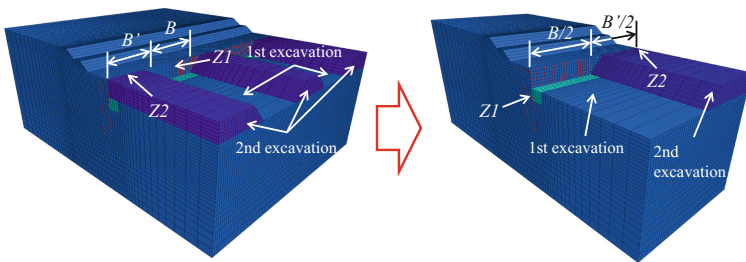


Fig. 7. Three-dimensional calculation model

4.1 Effect of Excavation Sequence

The construction sequences of non-horizontal struts foundation are usually as below: excavation, cushion, raft, and upper structure. For the foundations constructed by sections, the cushion and raft could be constructed when each section is excavated, or the sections could be excavated continuously. Therefore, three section excavation schemes are compared in this paper, as shown in Table 3.

Table 3. Three section excavation schemes.

Sequence	1	2	3	4
Scheme				
KW1	1 st excavation	2 nd excavation	cushion together	raft together
KW2	1 st excavation	1 st cushion	2 nd excavation	2 nd cushion
KW3	1 st excavation	1 st cushion and raft	2 nd excavation	2 nd excavation and raft

In the calculation model, the cushion is C20 concrete with 200 mm thickness, and the raft is C30 reinforced concrete with 1000 mm thickness. The section width is set as 60 m as an example. The front pile displacements of three different excavation sequences are shown in Fig. 8. In KW1, as the cushion and raft is not constructed until the 2nd section is excavated, the foundation is exposed for a long time, the displacements of 1st section are relatively larger than the 2nd section. The minimum displacement of piles is Z2.

The displacements of piles in KW2 and KW3 are both less than the scheme without section excavation. It can say that excavation by sections can significantly reduce the foundation deformation. Meanwhile, the displacement of Z1 is a little larger than Z2 in the scheme of KW2 and KW3, which means the cushion and raft can lower down the excavation impact if the cushion and raft is constructed. The difference from KW1 is that the cushion and raft play a role as the pit corner in KW2 and KW3, so the minimum pile displacement is at the interface of the two excavation stage.

Compared KW2 and KW3, the difference of the pile deformation at different places varies little, the restraining effect to the piles from the cushion is just a little lower than the raft. Therefore, the cushion construction can already greatly reduce deformation in non-horizontal struts foundation pit. In real projects, because of the engineering time limits, even if the KW2 or KW3 is adopted, the excavation at 2nd stage can hardly wait until the cushion and raft totally reaches designed strength, so the actual deformation is larger than the theoretical calculation results. For example, in Bailonggang wastewater treatment plant excavation, the cushion in A1 is almost constructed together with the excavation of A6 section, as shown in Figs. 3 and 4. However, the displacement of pile top is still smaller than the scheme without sections.

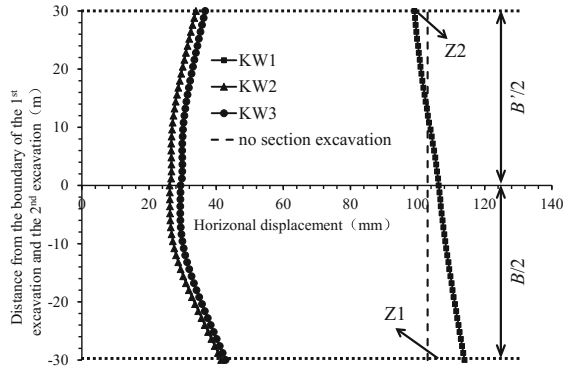


Fig. 8. Displacement of pile top by different excavation sequences

4.2 Effect of Section Width

In Fig. 9, the (a), (b), (c) show the horizontal displacements of pile top respectively at different width in KW1, KW2 and KW3. In different excavation sequences, the smaller the section width is, and the smaller the pile top displacement is. In KW1 scheme, the deformation can be rarely reduced even if the section width is very small. As shown in the curves, when the section width is 60 m, the displacement in 2nd excavation stage is even a little larger than that if the section width is set as 72 m. It may be caused by beam tension among piles. And the tension is larger if the section width is smaller. It can explain that the displacement of 60 m-wide section is relatively larger. If the section width is further reduced, the spatial effect of excavation is more obvious, and consequently, the displacement is reduced.

For KW2 and KW3, the pile top displacement is smaller when the section width is reduced. The biggest difference is located in the middle of 1st and 2nd excavation section, which is near Z1 and Z2. As the section width is reduced, the differences of Z1 and Z2 from the interface of 2 excavation stages both decreases. When the excavation width is smaller than 48 m, the displacements of different piles are basically the same. Continuously reducing the section width helps little for lowering the pile displacements. Therefore, there is an optimal section width for KW2 and KW3, when the displacements of Z1 and Z2 are close. For the condition in this paper, the optimal section width is about 24 m–48 m.

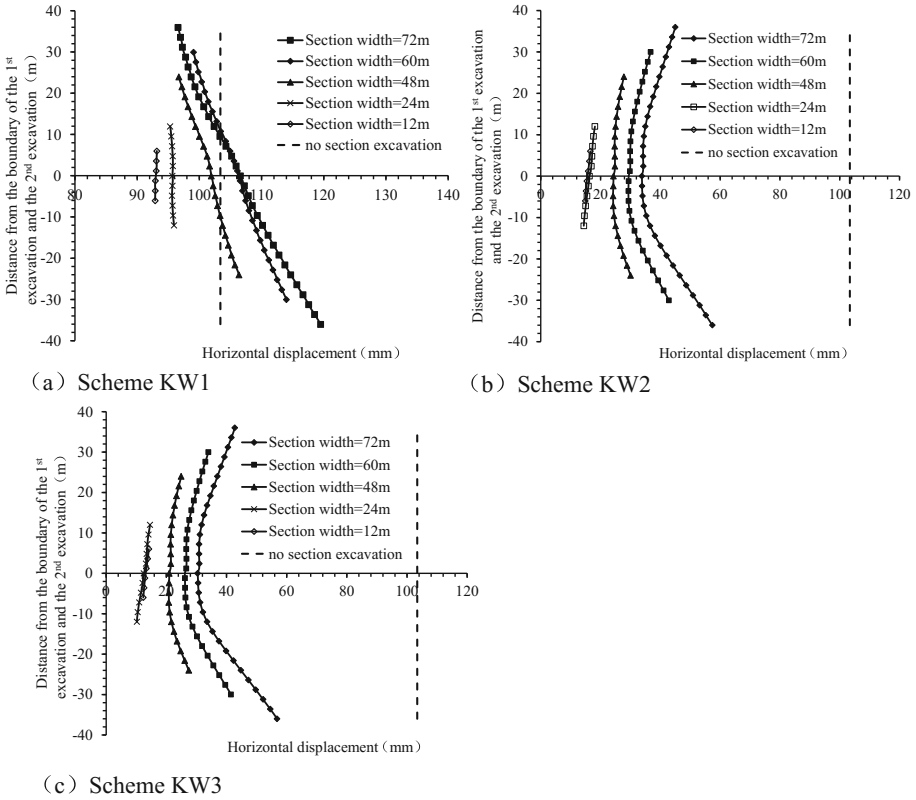


Fig. 9. The displacements of pile top at different section width

5 Conclusion

Based on the Bailonggang wastewater treatment plant foundation pit, the monitoring data are analyzed in this paper. And in addition, a finite difference calculation model is built, by changing the construction sequences and section width, three excavation modes are studied. The conclusions and suggestions are as below:

- (1) In Bailonggang wastewater treatment plant, the foundation without large area horizontal struts is excavated by sections. The adjacent soil can provide significant restraining effects for the piles of 1st excavation section, and meanwhile, the excavation of the 2nd excavation section has great impacts on the piles in the 1st excavation Sect.
- (2) When section excavation is adopted, the cushion construction as soon as possible can help to control pile deformation effectively. The effect of cushion is just a little lower than the raft, so the raft can be constructed when the 2nd excavation is finished.

- (3) The deformation of piles is reduced when the section width is smaller. For each excavation section, there is an optimal section width in the scheme of cushion or raft first. In the optimal condition, the displacements of piles at different places are close, additionally reducing section width cannot significantly lower down the pile deformation.
- (4) In this paper, only two excavation section with the same width is analyzed, but the foundation could be excavated by different widths or different section shapes in real projects. Therefore, further research of different excavation section types is still needed.

References

1. Liu, G.B., Wang, W.D.: Excavation Engineering Manual, pp. 94–95. China building industry press, Beijing (2009)
2. Ou, C.Y., Chiou, D.C., Wu, T.S.: Three-dimensional finite element analysis of deep excavations. *J. Geotech. Eng.* **122**(5), 337–345 (1996)
3. Lee, F.H., Yong, K.Y., Quan, K.C.N., et al.: Effect of corners in strutted excavations: field monitoring and case histories. *J. Geotech. Geoenviron. Eng.* **124**(4), 339–349 (1998)
4. Yu, J.L., Gong, X.N.: Spatial behavior analysis of deep excavation. *Chin. J. Geotech. Eng.* **21**(01), 24–28 (1999)
5. Gao, W.H., Yang, L.D., Shen, P.S.: Analysis of factors in time-space effect of force and deformation or retaining structure of deep foundation pit under soft soil. *China Civil Eng. J.* **34**(05), 90–96 (2001)
6. Lu, P.Y., Li, S.Z., Gu, L.Y.: Space analysis of retaining structures for foundation pits. *Rock Soil Mech.* **25**(1), 121–124 (2004)
7. Lai, G.Z., Fang, Y.G., Shi, H.Y.: Spatial mutual deformation analysis method for row of piles of deep excavation. *Rock Soil Mech.* **28**(8), 1749–1752 (2007)
8. Lei, M.F., Peng, L.M., Shi, C.H., An, Y.L.: Research on construction spatial effects in large-long-deep foundation pit. *Rock Soil Mech.* **31**(05), 1579–1584+1596 (2010)
9. Liu, N.W., Gong, X.N., Yu, F., Fang, K.: Analysis of spatial effects in strutted excavation and related influential factors. *Rock Soil Mech.* **35**(08), 2293–2298+2306 (2014)
10. Li, J.P., Chen, H.H., Li, L., Ma, J.S.: Observation on depth and spatial effects of deep excavation in soft clay. *China J. Highw. Transp.* **31**(02), 208–217 (2018)
11. Cheng, X.S., Zheng, G., Deng, C.H., Huang, T.M., Nie, D.Q.: Mechanism of progressive collapse induced by partial failure of cantilever contiguous retaining piles. *Chin. J. Geotech. Eng.* **37**(7), 1249–1263 (2015)
12. Cheng, X.S., Zheng, G., Diao, Y., et al.: Experimental study of the progressive collapse mechanism of excavations retained by cantilever piles. *Can. Geotech. J.* **54**(4), 574–587 (2016)
13. Do, N.A., Dias, D., Oreste, P., et al.: Three-dimensional numerical simulation for mechanized tunnelling in soft ground: the influence of the joint pattern. *Acta Geotech.* **9**(4), 673–694 (2014)
14. Zheng, G., Nie, D., Diao, Y., et al.: Numerical and experimental study of multi-bench retained excavations. *Geomech. Eng.* **13**(5), 715–742 (2017)
15. Do, N.A., Dias, D., Oreste, P.P., et al.: 3D modelling for mechanized tunnelling in soft ground-Influence of the constitutive model. *Am. J. Appl. Sci.* **10**(8), 863–875 (2013)



Research on Robot Education Curriculum System and Settings in Internet of Things Major

Ying Guo¹(✉), Lianzhen Zheng¹, and Yinong Chen²

¹ School of Information Science and Technology,
Qingdao University of Science and Technology, Qingdao, China
guoying@qust.edu.cn

² School of Computing, Informatics and Decision Systems Engineering,
Arizona State University, Tempe, USA

Abstract. Based on the characteristics of the existing Internet of Things (IoT) major, combined with the requirements of professional certification, this paper proposes to incorporate the robot education curriculum into the major course system that meets the graduation requirements and engineering curriculum. This paper presents the advanced robot education curriculum abroad, and proposes the teaching model of robot theory, robot experiment and robot comprehensive training. At the same time, the semester setting, the teaching content, the teaching hours, the credits and other related contents of the robot education course of Internet of Things major are also discussed.

Keywords: Internet of Things major · Robot education · Curriculum system · Course settings

1 Backgrounds

With the development of artificial intelligence technology, robots have acquired more and more attention. They play an important role in agricultural production, industrial manufacturing, and social services [1]. Countries around the world have started relevant research works [2], and China attaches great importance to robot technology innovation and industrial development [3, 4]. Training relevant talents who have a good knowledge of robots has become a significant part of modern education.

This paper studies the foreign robot education curriculum, which starts robot education from the middle school and high school. They improve the students' practical ability through cooperation and competition [5, 6]. Through the deeper and systematic study at the university, the students are more enriched and practical robot-related knowledge, possessing the ability to solve problems and innovate independently [7].

Drawing lessons from advanced curricula [8, 9], combining with the requirements of professional certification [10–14], and the characteristics of the existing Internet of Things specialty [14–18], we can take the robot education course as the teaching content of the Internet of Things major and incorporate it into the syllabus, which belongs to the professional certification of “professional courses that meet the

graduation requirements of this specialty”. In the process of teaching, the combination of robot theory teaching, robot experiment teaching and robot comprehensive training makes students receive more systematic robot education. At the same time, this paper discusses the courses of robot education, such as term, teaching content, teaching hours and credits.

2 Robot Education Curriculum System

Taking Arizona State University as an example, we discuss the teaching system of robot education course. Arizona State University has been ranked among the top universities for many years in the United States and its computer science ranks 32 worldwide. Its robot education program has been developed for many years with rich experience. Before the students enter the university, they will provide an introductory course for middle school students and high school students about robot education through summer camps, etc. In robot course learning, the students’ enthusiasm and creativity have been stimulated by competition, laying a good foundation for their study at the university level.

The robot education training for middle school students mainly involves the following contents:

- (1) 3D movie and game programming using Alice.
- (2) Draw 3D objects using Google SketchUp.
- (3) Create and deliver PowerPoint Presentation.
- (4) Construction of a LEGO EV3 robot.
- (5) Robotics programming in EV3 Graphical programming language.
- (6) First Lego League (FLL) competition preparation.
- (7) Program the EV3 robot with Bluetooth-remote control.
- (8) Program the EV3 robot to find its way out a maze with artificial intelligence.
- (9) Participate in the Robotics Competition at the end of the program.

Robot education training for high school students is based on the middle school and uses a more advanced programming environment. It mainly involves the following contents:

- (1) Basics of computing and logics behind computing.
- (2) Career path in science, engineering and computing.
- (3) Advanced PowerPoint presentation development.
- (4) Visual IoT/Robotics Programming Language Environment.
- (5) Program robot in Virtual environment and in physical environment.
- (6) Program the robot with Bluetooth-remote/Wi-Fi control.
- (7) Program the robot to find its way out a maze with artificial intelligence.
- (8) Participate in the Robotics Competition at the end of the program.
- (9) Web programming and Web app development.
- (10) Smart phone programming, Android Phone App development and Windows Phone App development.

The training involves not only the software design, but also the assembly of the robot hardware. Students complete the design through group cooperation and display their design results in the form of competition.

At the university level, there are many programming courses. For example, “Introduction to Programming Languages” related to robot software is opened twice a week, 75 min one time, “Distributed Software Development” is in the same way, “Introduction to Engineering” about robot hardware and corresponding experiments is opened once per week, 50–170 min each time. There are a variety of robot-related elective courses (Alice, EV3) and competition training. Students learn and practice via more systematic and professional courses, and have a good knowledge of robot software design, hardware assembly, design and development.

3 IoT Major Robot Education Curriculum

Based on the advanced robot education curriculum and teaching methods, taking the Internet of Things major of Qingdao University of Science and Technology in China is as an example to discuss the settings of robot education courses. There is no robot education course in middle schools and high schools in China. The settings of robot education courses should be from breadth to depth, and students’ knowledge systems and ability should be gradually established.

3.1 Course Semester

The core courses of Internet of Things major of Qingdao University of Science and Technology include Introduction to Internet of Things, Fundamentals of Electronic Technology, Java Programming and Application, Control Foundation, Internet of Things Communication Technology, Radio Frequency Identification Technology, Wireless Sensor Network, Embedded System, Artificial Intelligence, Network, Information Security, cloud computing, data mining, etc., does not include robot education courses.

Based on the Control Foundation and Artificial Intelligence courses in the Internet of Things major, robot education courses can be added to the existing curriculum system, as shown in Fig. 1.

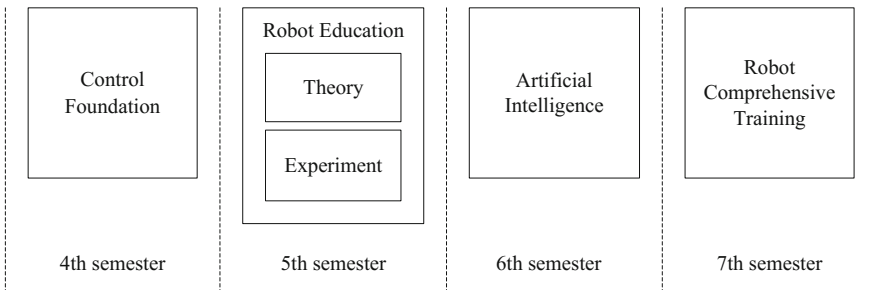


Fig. 1. Robot education courses and corresponding semesters

The robot education courses can be divided into three parts. The first part is theoretical study, the second part is experimental study, and the third part is training. The theoretical study course can be opened in parallel with the experimental study course, arranged after the Control Foundation course, before the Artificial Intelligence course. The training course can be set up after the Artificial Intelligence course, allowing students to bring the Control Foundation, Robot, and Artificial Intelligence courses together and understand thoroughly. The training course can improve students' manipulative and design ability.

The Control Foundation and Artificial Intelligence courses of Internet of Things major of Qingdao University of Science and Technology are offered in the 4th and 6th semesters respectively. The Robot Education courses can be established in the 5th semester, and the comprehensive training of robot can be opened in the 7th semester, forming a complete knowledge system.

3.2 Robot Education Course Contents and Hours

3.2.1 Robotics Theory Course

The key technology used in the first-year robotics course is the use of the Visual IoT/Robotics Programming Language Environment (VIPLE) [19–21]. It applies computational thinking idea by pre-programming the major components into services and by integrating the components through drag-and-drop composition. It makes it easier for students focusing on the computational logic, instead of syntax and programming details.

The course mainly covers the basic knowledge of robots, software design, hardware design and other aspects. It can set 24–32 h, 1.5–2 credits. The course contents and time schedules are as follows (Table 1):

Table 1. Robotics theory course contents and time schedules

Teaching contents	Class hours
Engineering design process	1 class hour
Robot overview	1–2 class hours
Robot development environment and VIPLE	4–6 class hours
Logic design and computer composition	3–4 class hours
Event-driven programming and finite state machines	3–4 class hours
Robot hardware composition	5–6 class hours
Service computing and web application development	4–5 class hours
App development on android mobile	3–5 class hours

The robot education theory course can be divided into 8 parts according to the teaching contents: Engineering Design Process. The engineering design process consists of the following steps:

- Identify and define problem
- Literature review and research
- Sketch possible solutions
- Modeling
- Analysis
- Simulation
- Prototyping
- Final selection
- Implementation and testing
- The engineering design process is used as the guideline for the rest of the components. The following diagram illustrates the process (Fig. 2).

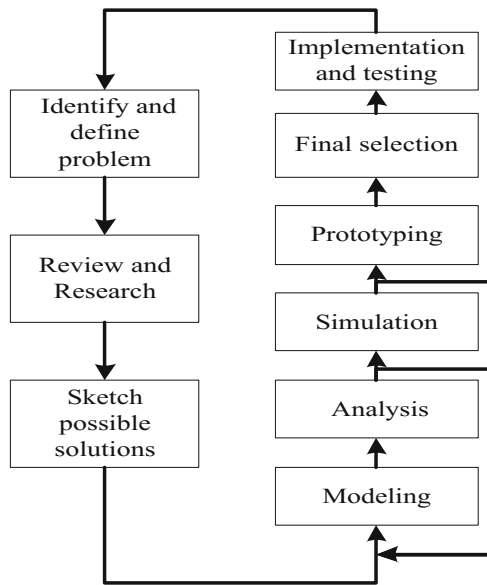


Fig. 2. The engineering design process

Modeling, analysis, and simulation will be the quantitative part of the engineering design process and will be thoroughly exercised through finite state machine and various simulation tools used in the course.

Robot Overview. Since there are no robot education courses in middle school and high school universally in China, the overview part is an introduction to robot development technology, which gives students a preliminary understanding of the course contents, the basic knowledge they should have, and the learning objectives to be achieved.

Robot Development Environment and VIPLE. The Robot Development Environment VIPLE is a visualized networking/robot programming language environment developed by the Arizona State University’s IoT and Robotics Education Lab. It has

open API and interfaces to support multiple IoT and robot platforms. This section focuses on workflow and visual programming, the VIPLE IoT/robot development environment, the engineering design process of VIPLE, the activities and services of VIPLE, the use of VIPLE (use or with If activities to create conditional loops, use While loops, use global variables to create events, create Counter activities, build 2-1 multiplexers, etc.).

Logic Design and Computer Composition. This part involves much knowledge about computer, and can add or delete learning contents according to the students' learning situation of logic circuit and computer composition. It mainly describes the type composition of computer systems, and creates computer system components in VIPLE, such as: creating logical AND gates, creating 1-bit full adders, creating 2-1 multiplexers, creating 4-1 multiplexers, creating 1 bit ALU, automatic test, etc.

Event-Driven Programming and Finite State Machines. Event-driven programming and finite state machines are the core components. They mainly teach event-driven programming finite state machines, use ASU VIPLE to solve event-driven problems, event-driven counters, parity checking, parallel computing, and wire-controlled simulation, etc.

Robot Hardware Components. Different contents can be taught depending on the hardware conditions of the school. Mainly involving calculation and communication model in VIPLE, robot hardware overall structure, main control board, Intel Galileo development board, Intel Edison module, Arduino/Genuino, TI CC3200 LaunchPad, Raspberry Pi for robot master module, sensor module (ultrasonic sensor/infrared sensor/light sensor/color sensor), steering gear, Galileo robot, Edison robot hardware and software installation, etc.

Service Computing and Web Application Development. Service computing and web application development include parallel processing technology, basic concepts of text language programming, conceptual framework of service-oriented architecture, Visual Studio programming environment, parallel and service-oriented computing in VIPLE, and secure applications adopting encryption/decryption services, etc.

App Development on Android Mobile. This section mainly introduces the application of App Inventor on Android phones, which can be combined with the experimental part.

3.2.2 Robot Education Experiment Course

Through experiments, students can put the theoretical knowledge of robot into practical operations, and control robots to achieve various functions. It includes 4 experiments; each of them takes 4–8 class hours, total of 16–32 class hours, 0.5–1 credits. According to the hardware equipment of the school, it could increase or decrease the experiment time, the contents and time schedules are as follows (Table 2):

Table 2. Robot Education Experiment Course Contents and Class Hours Schedule

Teaching contents	Class hours
Robots in the simulated environment and maze navigation	4–8 class hours
Intel/Raspberry Pi robot programming	4–8 class hours
LEGO EV3 robot programming	4–8 class hours
Android mobile programming	4–8 class hours

The main points involved in the experiment courses and the contents that students need to master are as follows:

Robots in the Simulated Environment and Maze Navigation. This is a software simulation experiment that can be performed even without hardware. Students should familiar with the VIPLE machine services and the robotic platforms supported by VIPLE.

Design and implement an algorithm that traverses the maze, use a finite state machine of maze navigation algorithms, and complete an autonomous maze navigation algorithm in the VIPLE simulator. Understand the maze algorithms, along the wall algorithms, program the web robot to make it go around the right wall, use the two-distance local optimal algorithms to traverse the maze; understand the unity simulator and VIPLE program, implement VIPLE block diagram, design the activities of the two-distance local optimal algorithms and the main block diagrams.

Familiar with Web 2D simulators and Web 3D simulators, configure VIPLE to use Web simulators, implement Main block diagram of the wall following algorithm in Web simulators, implement activities related to the wall following algorithm in Web simulators, the main block diagram of the two-distance local optimal algorithms in Web simulators, the involving activities of the two-distance local optimal algorithms implemented in the Web 2D simulators.

Intel Robot Programming. For Intel robots, use the maze navigation employing the wall following algorithm to implement Init activities, Left1 and Right1 activities, Right90 and Left90 activities, Backward and Forward activities, and ResetState activities.

The maze navigation using the local optimal algorithm solves the maze problem with the two-distance algorithm, controls the Intel robots in the VIPLE, and implements the activity in the VIPLE programs.

Use the finite state machine; use event-driven programming for maze navigation, using event-driven programming of the Left90, Left1 and Backward activities. Familiar with the Main block diagram based on event-driven activities, use a light sensor to implement a basic sumo algorithm.

LEGO EV3 Robot Programming. For LEGO EV3 robots, get sensor readings from EV3 Brick, bluetooth connection, acquire sensor readings through the program, connect robots to VIPLE via bluetooth or Wi-Fi, remotely control the EV3 robots, and drive robots by means of connection in VIPLE to improve the drive experience.

Use VIPLE's tracking and sumo robot programs, install color sensors, track, use the light sensors to implement the basic sumo algorithms, use the light sensors' and the touch sensors' sumo algorithms.

Employ EV3 wall following program of VIPLE, navigate along the wall maze (Main block diagram), configure sensors in the wall following algorithm, use wall following program adopting event-driven programming, make use of the local optimal heuristic algorithm for labyrinth navigation, and implement the local optimal algorithm of Main block diagram.

Android Mobile Programming. Android mobile programming allows students to install APP programs on their mobile phones, implement Hello World programs, Moore mud games, stock quotes, stock trends, plate shooting games, shoot multiple plates, play brick games, and apply App Inventor to program NXT Robot, number guessing games, simple side skateboarding games, memory games, etc., under the guidance of teachers.

3.2.3 Robot Training Course

According to the existing hardware resources of the universities and the students' mastery of artificial intelligence algorithms, students are grouped by the instructors depending on the students' abilities. Different tasks are arranged separately, and completed gradually. The course can be arranged for 1-2 weeks. Moreover, Robot Education integrating training and Robot Competition can be combined to encourage students participate more. They could solve problems through teamwork, and improve their capabilities of design and application.

4 Conclusions

Incorporating the current international advanced Internet of Things and robot education platforms and methods into the domestic IoT professional teaching system and opening it as a major course for all students cannot be done overnight. It needs to be gradually developed according to the professional specialties, hardware conditions and students' quality of different schools. The goals are accumulating experience in practice, constantly improving the teaching contents, and continuously improving the teaching methods to achieve the desired teaching effects.

Acknowledgments. The project is supported by Natural Science Foundation of Shandong Province (Number ZR2016FQ10).

References

1. Huang, R., Liu, D., Xu, J., Chen, N., Fan, L., Zeng, H.: Development status and trend of educational robots. *Now Educ. Technol.* **27**(1), 13–20 (2017)
2. Wang, J., Hu, L., An, L.: Case study of foreign robotic STEM educational robot course. *Now Educ. Technol.* **27**(4), 13–20 (2017)

3. Ge, Y.: Research on Key Technologies of Educational Robots Based on Internet of Things. Wuhan University of Technology, Wuhan (2013)
4. Ge, Y., Li, W., Zhou, J.: Thoughts on introducing IOT education robot into undergraduate teaching. *Educ. Teach. Forum* **2016**(03), 203–204 (2016)
5. Xu, D.: Research on Knowledge Mapping of Robot Education. Bo Hai university, Jinzhou (2018)
6. Chen, Y., De Luca, G.: VIPLE: visual IoT/robotics programming language environment for computer science education. In: *IPDPS Workshops*, pp. 963–971 (2016)
7. Lei, L., Liang, M.: Visual analysis of domestic robot teaching research. *Softw. Guide (Educational Technology)* **15**(08), 75–77 (2016)
8. Chen, Y., Zhou, Z.: Service-oriented computing and software integration in computing curriculum. In: *IPDPS Workshops*, pp. 1091–1098 (2014)
9. Chen, Y., Zhou, Z.: Robot as a service in computing curriculum. In: *ISADS*, pp. 156–161 (2015)
10. Guo, Y.: Professional certification-oriented courses for Internet of Things engineering. *Internet Things Technol.* **87**(05), 115–117 (2018)
11. Wang, H., Liu, C.: Integration of IoT engineering education resources based on teaching route from the perspective of professional certification. *J. Nantong Vocat. Univ.* **31**(1), 27–32 (2017)
12. Xu, D., Zuo, X., Teng, W.: Research on talents training model of Internet of Things engineering based on professional certification. *Technol. Vis.* **2015**(35), 50–54 (2015)
13. Wu, S.: Exploration and reform of practical teaching of Internet of Things engineering under the background of engineering education certification. *Internet Things Technol.* **7**(11), 109–111 (2017)
14. Guo, Y., Zeng, X.: Research on the theory and technology curriculum of IoT professional control. *Comput. Educ.* **2013**(16), 86–89 (2013)
15. Min, T., Yong, S., Xi, L.: Exploration and construction of the course system of Internet of Things application technology in higher vocational colleges. *Internet Things Technol.* **4**(2), 85–87 (2014)
16. Guo, Y.: Research and exploration of tiny OS course teaching in Internet of Things. *Comput. Educ.* **2015**(14), 86–89 (2015)
17. Wang, Z., Zeng, X., Wang, X.: Construction of teaching materials for compulsory courses of Internet of Things Engineering. *Comput. Educ.* **2012**(21), 5–8 (2012)
18. Ma, F.: Exploration on the Introduction of Internet of Things Professional Teaching in Robotics Course
19. *Information and Communication* **2017**(08), 23–24 (2017)
20. Chen, Y.: *Service-Oriented Computing and System Integration*. 6th edn. Kendall Hunt Publishing (2018)
21. Chen, Y., Chen, W., Han, D.: *Introduction to Computer Science and Engineering: Based on IoT and Robotic Visual Programming Practice Methods*. Mechanical Industry Press (2017)
22. Chen, Y., Cai, W.: *Service-Oriented Computing and Web Data Management-From Principles to Development*. Xidian University Press, Xi'an (2013)



Research on Expressway Network Tolling Platform Technology Based on GIS+BIM

Li-xia Bao¹(✉) and Yuan-qing Wang²

¹ Shanghai Urban Construction Design and Research Institute (Group) Co., Ltd.,
Shanghai 200125, China

jlublx@163.com

² School of Automotive and Transportation Engineering,
Liaoning University of Technology, Jinzhou 121000, China

Abstract. Architecture of expressway network tolling platform is put forward based on GIS+BIM, real-time tolling fee calculation for each vehicle in networked and prewarning of equipment and facilities as well as data analysis and mining are studied, the basic network diagram of GIS+BIM is presented, and the database and data dictionary are researched. Platform performance requirements such as high efficiency of calculation of fee, transparency of business management, data identification, high fault tolerances of storage, and rapid response are studied. The research of this paper will provide certain reference for the construction of expressway tolling system.

Keywords: Expressway network tolling platform · GIS+BIM basic network · Charging data dictionary

1 Introduction

The expressway is an important transportation hub in China and plays an important role in promoting economic development. With the rapid improvement of people's material life, the way of travelling is more and more diversified. The traffic volume of expressway tolling stations is also increasing and the tolling system is continuously built and improved. The modernization construction of the tolling system is gradually realized, which is invincible in today's highly competitive transportation industry [1]. The charging method is also changed by the original and independent running of MTC to the current operation based on MTC+ETC, and combines with the application of the intelligent expressway tolling platform's architecture supported by scientific and technological information such as computer information technology, network technology, communication technology and sensor technology, which can speed up the tolling efficiency of the expressway, effectively improve the operational efficiency of the expressway and thus ensure its smoothness.

2 Architecture of Expressway Network Tolling Platform

With the increase of expressway network mileage and the development of technology, the functions and performance requirements of the expressway tolling system platform are becoming increasingly high in the country and various provinces. The domestic and foreign experts and scholars have carried out the research on charging methods. From 1990 to the end of 1995, European countries mainly implemented the second-generation electronic tolling collection system, which adopted the 5.8 GHz DSRC frequent and readable/writable electronic tags recommended by the EU [2]. In 2001, Japan began actively using ETC (electronic tolling collection) technology. Until December 2012, more than 52 million vehicles had been installed ETC's dedicated on-board units, accounting for 70% of the total number of vehicles [3]. In 2005, the German expressway adopted an electronic charging method based on the combination of mobile communication and GPS satellite positioning system [4]. In China, Wang [5] proposed the use of 5.8 GHz dedicated short-range wireless communication technology to complete the data communication between RSU and OBU through the analysis of ETC system. Zhang [6] used the architecture of the B/S to design the expressway network charging inspection management system, which is uniformly deployed by the network center. The charging inspection system can realize electronic rapid processing, and improve the work efficiency and the inspection service level. Li [7] proposed the design concept of the charging system based on "Internet+", and introduced a third-party payment platform to make it have the advantages of high service efficiency and low implementation cost. The above research only designed a single system to solve a single problem, and did not link the systems together, making them unified networking, layered design, and synergy.

This paper combines the existing business needs and builds the architecture of the expressway network tolling platform, which mainly includes 7 major systems and 4 major platforms. The mutual connection is shown in Fig. 1. Among them, the comprehensive functions of expressway tolling collection system will mainly be realized such as the data aggregation of the tolling lane and the identification station, the issuance of the rate table, the calculation of the ambiguous path's fee, the reconciliation, the clearing, the splitting, and the docking with Ministry of Communications Road Network Command Center. The data of the MTC and ETC's tolling lanes is transmitted to the expressway tolling clearing and settlement center, where the portal data aggregation, the ambiguous path identification and the calculation of fee, export charging management, clearing settlement, auditing, statistics, sharing and charging as well as big data analysis are performed on the platform. The platform also provides independent authority for each section center and tolling station. The operator can query the charging and related data within the authority in real time. The card company transmits and shares data through the dedicated line and the clearing center. The clearing center exchanges data with the national center through the private network. The third-party payment platform settles with the expressway clearing and settlement center through the Internet and the system reserves the expansion business of ETC in the future. The social parking garage is docked with the expressway safety management system. The parking space will be provided to avoid the vehicles occupying the emergency lane due to abnormal events and facilitating the smooth flow of vehicles in need.

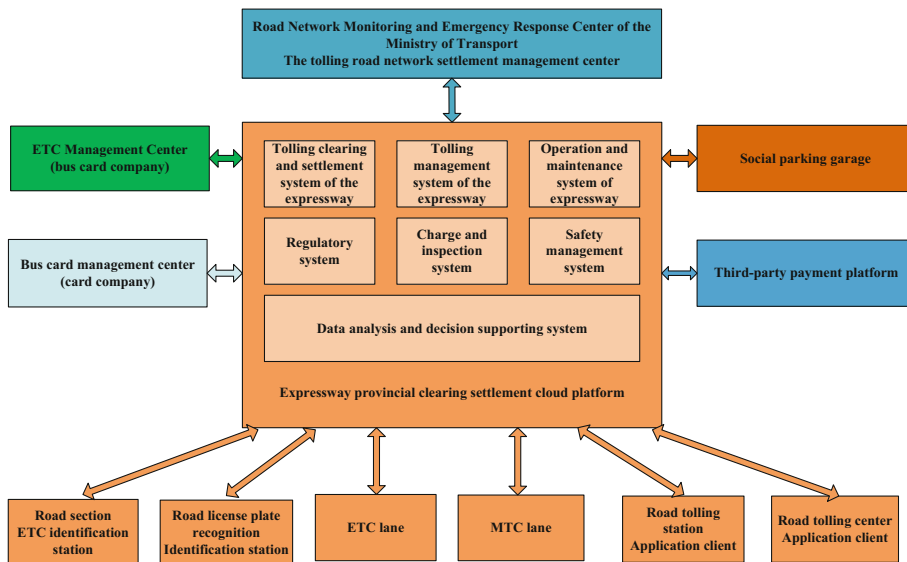


Fig. 1. Three-level architecture diagram of the charging system

As shown in Fig. 1. The core subsystem of the charging software system mainly includes:

(1) **Subsystem of the clearing and settlement center**

The subsystem of the clearing and settlement center mainly consists of some modules such as the GIS layer and basic data management of the tolling network, the data aggregation of the entrance and exit and the identification station, the ambiguous path identification, the charging management, the tolling splitting and settlement, aggregating and sharing of the data, and the operation and maintenance, auditing, supervision, data analysis and decision supporting, blacklist update, rate update, client function of each road segment, tolling station client function, systematic parameter management, report statistics/inquiry/printing, ticket and traffic ticket management, data transmission, data storage, data backup and recovery, information releasing, network management, operation authority management and key management as well as security authentication. Each module cooperates to realize the refined management of multi-aspect information perception, aggregation, transmission, sharing, supervision, real-time update and other functions.

(2) **Subsystem of the link center and tolling station management**

The charging software system is based on the architecture of the cloud computing Pass, which realizes the common functions of each road segment center and tolling station, and these functions are all realized in the form of services in the clearing and settlement center subsystem using the platform. The road center and the tolling station are based on the client to access clearing and settlement subsystem, and perform the functions

such as the tolling collector's commuting assignment, the dispatching and recycling request of the password card, the reconciliation, the auditing, the clearing query, and the report within the scope of the road section. For the special needs and special functions of certain road sections, it can be developed according to the needs in the subsystem of clearing and settlement center.

(3) Subsystem of the tolling lane

The tolling lane's subsystem is divided according to the MTC and ETC lanes. In order to meet the different needs of users and speed up the tolling efficiency of expressways. The subsystem of ETC lane mainly includes modules such as RSU management, receiving and updating center system's operating parameters, real-time fee receiving, automatic deduction, lane equipment management, local fee calculation, and special handling. The MTC lane subsystem mainly includes modules for issuing cards, receiving update center system's operating parameters, license plate's identification information aggregation, real-time fee receiving, charging, clearing card, lane equipment management, local fee calculation, and special handling. To realize the charging management's block processing.

The architecture of the expressway network tolling platform uses the above subsystems to realize the functions of ambiguous path identification, real-time management of vehicle tolling collection, the auditing, data aggregation, data transmission, data storage and security sharing. In view of the above functions, the expressway tolling platform based on GIS+BIM realizes the real-time calculation of networked fee and the unified integration of central tolling management of each section, and develops the charging network base layer to record the characteristic data of each section in real time and establishes a BIM model, which can display tolling equipment facility's characteristic data and operational status information in real time.

3 Key Technology of Expressway Tolling Platform Based on GIS+BIM

3.1 Functional Goal

The GIS+BIM platform and the big data platform are mainly developed on the expressway clearing and settlement software system, which will deploy the existing fee-based business application system and the newly developed precise fee calculation, charging inspection, asset management to put in the new basic platform, to provide real-time calculation of networked fees, management of various sections of tolling stations and data analysis and decision-making. Developing expressway network GIS platform based on ArcGIS or Supermap, which newly develops tolling network foundation layer. Adopting two dimensional GIS on road section can record the characteristic data of station number, number of lanes, completion time, affiliated section company, rate, etc. The entrance and exit lane and station adopt the BIM model to display railings, ceiling lights, tolling booths, readers, license plate identification and other facilities on the tolling lanes, and form a basic road network geographic information database. All operated fee, charging, clearing, statistics, auditing, assets

management and so on are based on GIS+BIM networks. But the platform center needs to calculate the amount of fees which already incurred by the connected vehicles in real time and distributes them to the possible two-tier exit tolling lanes. When the fee is audited, it will need to adopt a certain image analysis and learning function. Therefore, it is necessary to design and construct a big data base platform to meet fast calculations of the massive data.

The basic platform of expressway tolling system based on GIS+BIM mainly achieves the following goals:

- (1) Establishing a GIS+BIM platform to manage the characteristic data of road sections, tolling lanes, and signage stations. It has the functions of adding and modifying tolling road sections, lane types, signage station sites and equipment facilities.
- (2) Relying on GIS to realize all dynamic traffic data, ambiguous path identification, fee calculation, and charging dynamic data's fee in any time range and any location.
- (3) Establishing a BIM model in the tolling station, each lane and signage station and integrating with the GIS network to load the tolling lane railings, ceiling lights, fee-card displaying, license plate identification, RSU and characteristic data of other equipment facility and display real-time operational status information.
- (4) The network cost function model is established on the basic platform, and the dynamic rate can be adjusted in real time. The accurate paths and fees of each vehicle are calculated in real time based on the network topology diagram, the aggregated entrance and exit lanes and the identification station. At last, the processing is delivered and downgraded according to the algorithm.
- (5) The data such as traffic volume, speed, and license plate of any segment in real time can be displayed on the basic platform.
- (6) The basic platform provides business application systems for employees' commuting registration, pass-through card issuance, bill issuance, tolling collection, charging clearing, inspection, equipment maintenance, etc.
- (7) Displaying the running status of the driveway and the identification station, alarming and repairing on the basic platform.
- (8) Launching the image analysis, charging big data analysis (total charging, road sections, tolling stations, etc.), blacklist analysis, special vehicle painting, charging audits, OD inference and other big data analysis and mining.

3.2 GIS+BIM Basic Network Diagram

(1) The basic network layer development based on GIS

The established GIS system covers the static characteristic information related to the tolling expressway, but it has not been associated with the tolling service system at present, and the data in terms of dynamic rate, asset data, and equipment status information is not complete enough. The platform's development is based on the existing characteristic data, and needs to be further checked. As shown in Fig. 2. Any two points of the tolling station, lane division, road segment, and identification station

are added to the GIS foundation layer of the expressway tolling network. The functions such as the interval amount layer, the affiliated road segment company, and the dynamic data will be displayed.



Fig. 2. Expressway tolling network GIS foundation layer

(2) BIM construction of tolling stations and identification stations

The BIM model has been widely used in the design, construction and operation of buildings and municipal engineering. The platform is developed to clearly display the tolling booth and 32 identification stations. This BIM model is established as shown in



Fig. 3. BIM model of the tolling station (diagram from <http://www.ddove.com>)

Fig. 3 and the lane width, type, and various equipment facilities are presented in a three-dimensional manner. The BIM model is combined with the GIS platform to generate an operation and maintenance platform, which can simulate released/closed statuses of the lane in real-time and collect data such as the tolling collector's ID, the passing information, the card issuing information, the charging, the OBU identification, the operating status of various devices, and the ticket. Based on the data of a single lane on this platform, big data statistics and analysis can be performed for any tolling station and any time.

3.3 Database and Data Dictionary

At present, the database used by the expressway tolling collection system is Oracle9i and MySQL. Because the definition of the data dictionary is not standard and the protocol of the interface is not uniform. It is increasingly unable to meet the increasing data management requirements of the traffic volume and the ambiguous path collection. The Network Center is developing a data dictionary for the National Networked Tolling Center. Therefore. According to the national requirements and relevant regulations, the charging system designs the data dictionary which is required to be compatible with the existing database in the database selection to ensure the smooth switching between the old and new systems. The definition of data in the platform should follow the key information coding and other relevant regulations in the National Expressway Network Naming and Numbering Rules (JTG-A03-2007), Tolling Road Networking Charging Ambiguous Path Identification Technical Requirements (2015). The data categories involved in the expressway tolling collection system mainly include:

- (1) Expressway network's GIS basic characteristic data (including 4 provincial tolling stations), road section, tolling station characteristic data definition, which includes the completing date, opening date, affiliated company, nature (government loan or business) and flexible charging data;
- (2) Including the basic data for charging, blacklist, the shortest path rate, vehicle model, dynamic rate between any two points and amount of fees, etc.;
- (3) The ambiguous path recognition model and calculation data;
- (4) License plate identification data: license plate data for entrances and exits and identification stations;
- (5) Clearing and settlement business data: including business data such as consumptions, statistics, accounting, splitting rules, clearing and settlement, and owner's income of Shanghai Road Network;
- (6) Modifying and increasing the billing data: including the invoiced statistics, each invoice mileage, the amount of fees, the splitting flow data;
- (7) Road segment business data: road segment's characteristic data, tolling flow and statistics of each road segment, accounting, etc.;
- (8) Tolling station business data: tolling station's characteristic data, the charging data, statistics, accounting and other business data of each tolling station;
- (9) The data of blacklist and auditing;
- (10) Various types of statistics, mining and other post-processing data;

- (11) Outfield equipment operation, monitoring and maintenance data: including operation status and maintenance data of equipment such as lanes, signage stations, communications, etc.;
- (12) The data of the clearing and settlement center server, storage, cloud platform, software and other operational monitoring and maintenance update data;
- (13) Interactive and shared database: including systems that related to relative data of other charging system

The data data required by the system platform is shown in Fig. 4. It mainly includes the charging clearing and settlement data, the charging inspection data, the operation and maintenance data, the charging management data and the changing data.

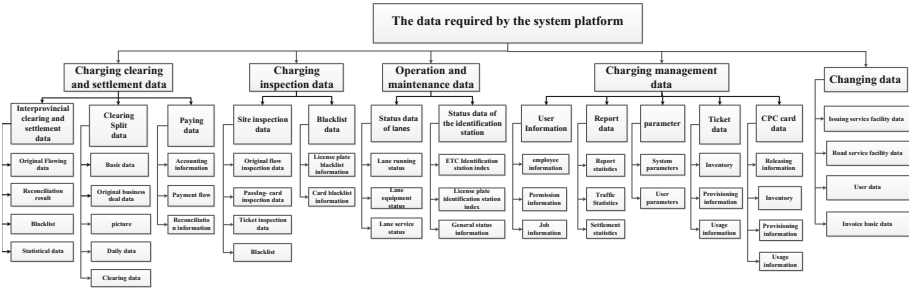


Fig. 4. Data required by the platform

4 The Research of the Platform Performance Requirements

The expressway clearing and settlement software system can achieve the following performance indicators under the joint action of developing GIS basic network layer, BIM construction of tolling station and identification station, database and data dictionary design.

- (1) The number of users that they have uplink real-time communication needs of clearing and settlement center and tolling lanes and signage stations is designed to be 3,400 times/s, and the maximum downlink communication demand is 10,000 times/s;
- (2) The central fee calculation requirement in real time is designed to be 3000 times/2 s;
- (3) The clearing and settlement center will provide the business center management (working entry, passing ticket management, charging inquiry, bill management, clearing inquiry, off-duty entry, report generation, etc.) within the field of the road segment center. The number of concurrent users is designed to be 30 and the response time requirement of page request is needed less than 1 s;
- (4) The clearing and settlement center shall handle more than 3 million transactions per day;

- (5) Under the normal conditions. Vehicle identification information, card issuance information and transaction data of MTC lanes and ETC lanes should be uploaded to the clearing and settlement center in real time and the delay is less than 1 s.
- (6) ETC trading time is less than 270 ms;
- (7) ETC lanes: the lifting time and landing time of the electric railing of the main line tolling gate shall not exceed 0.4 s, and the lifting time and landing time of the electric railing of the ramp tolling shall not exceed 0.6 s;
- (8) The number of transaction flow data that can be saved in the lane is more than 100,000 and the storage time is not less than 40 days;
- (9) Normal ETC vehicles pass the ETC lane at a speed of 20 km/h, and the error of per thousand electronic tolling transactions is no more than 3times;
- (10) License plate recognition record that can be saved in the lane (including one picture per record) is more than 100,000;
- (11) The 5.8G identification data of the identification station is synchronously written into the OBU and the CPC card, and the license plate identification data is uploaded to the clearing and settlement center and its delay is less than 2 s;
- (12) The 3D model of tolling station and signage station has a quick response time of less than 10 s.

5 Conclusion

The tolling station of the expressway is the node connecting the highway and the expressway. Its management level seriously affects the operational capability. However it simultaneously reflects the operation of the expressway. Architecture of expressway network tolling platform is put forward based on GIS+BIM, real-time tolling fee calculation for each vehicle in networked and prewarning of equipment and facilities as well as data analysis and mining are studied, the basic network diagram of GIS+BIM is presented, and the database and data dictionary are researched. Platform performance requirements such as high efficiency of calculation of fee, transparency of business management, data identification, high fault tolerances of storage, and rapid response are studied. Relying on the BIM model and the GIS platform, the operation and maintenance platform can be generated, the actual running charging status can be simulated, and the big data of any tolling station and time-period based on the single lane data on the platform can be calculate and analyze. The charging system is more efficient, accurate, reliable and stable, and thus improves the service level and operational capability of the expressway. It has certain reference for the construction of expressway tolling system.

References

1. Wang, Y.: Research on the innovation of the management system of expressway tolling stations. *Green Build. Mater.* **06**, 84 (2017)
2. Ma, W.: Multi-function non-stop charging system based on DSRC. *Comput. Eng. Appl.* **10**, 246–248 (2002)

3. Sun, W.-b.: Research on multi-path tolling splitting of expressway network tolling. Tianjin University (2014)
4. Yang, Y.: German expressway electronic tolling collection system. *Urban Traffic* **04**, 82–84 (2006)
5. Wang D.: Design and implementation of expressway electronic tolling collection system in Sichuan Province. University of Electronic Science and Technology of China (2015)
6. Zhang, R.-y.: Design and implementation of expressway network charging and inspection management system. Tianjin University (2016)
7. Li, R.: Analysis of “Internet+” Expressway Tolling System. Taiyuan University of Science and Technology (2017)



Commercial Vehicle Dynamic Third Party Safety Supervision of Based on GPS/Beidou Technologies

Guangyue Nian¹, Jinping Li²(✉), Daniel(Jian) Sun^{1,3}, and Xinhan Li⁴

¹ Center for ITS and UAV Application Research Center,
School of Naval Architecture, Ocean and Civil Engineering,
Shanghai Jiao Tong University, Shanghai 200240, China
ngyown@sjtu.edu.cn

² Beidou Information Institute of Technology,
Shanxi Beidou Weixun Electronic Technology Co., Ltd., Taiyuan 030009, China
railsj@sjtu.edu.cn

³ State Key Laboratory of Ocean Engineering, School of Naval Architecture,
Ocean and Civil Engineering, Shanghai Jiao Tong University,
Shanghai 200240, China

⁴ Chongqing Institute of Intelligent Transportation System and Technology,
Chongqing Mingxin Changxing Technology Co., Ltd.,
Chongqing 400020, China

Abstract. Based on GPS/Beidou technologies, dynamic safety supervision of commercial operating vehicles has played an important role in transportation industry safety. From the serious traffic accidents occurred in recent years, the major reasons were the improper safety supervision of vehicles. The traditional dynamic safety supervision mode and applications are insufficient. To make up for the lack of traditional dynamic supervision and realize the traceability and visualization of the supervision process, it is necessary to innovate the road transportation safety monitoring, with new management concepts and technologies. Based on the practical understanding of the traditional monitoring status of the industry, this paper proposes the third-party safety supervision of commercial vehicles. We successfully incorporate the application of third parties in other fields world widely and combine it with the development of vehicle dynamic monitoring technology. Through the innovative application of empirical applications in certain areas, the government bureau, enterprises, and monitoring operators all recognized the new model. The overall safety level indicators and accident data before and after the application are compared and analyzed. The results indicate that the new model has largely improved the safety level of transportation within the region. At the same time, it has broadened the scale of supervision, practically reducing the corporate costs and operating investment. The new model greatly improves the dynamic security monitoring effect, which satisfies all parties' demands without increasing the input. At present, this model has been promoted and applied in Xinjiang Uygur Autonomous Region, which was proved to have important social and economic significance for modern transportation industries.

Keywords: Third party · Safety supervision · Dynamic · GPS/Beidou · Transportation safety

1 Introduction

Safety production is an eternal theme of the road transportation industry and an important guarantee for the sustainable development of the road transportation industry. According to the statistics of the Ministry of Public Security, China, in 2016, there were a total of 50,400 road traffic accidents involving trucks, resulting in 25,000 deaths and 46,800 injuries, accounting for 30.5%, 48.23%, and 27.81% of the total number of motor vehicle accidents [1], which indicated the seriousness of commercial vehicle accidents. Commercial vehicles mainly include trucks and buses, which are prone to cause accidents such as mass casualties.

In 2014, three ministries and commissions promulgated the “Measures for the Supervision and Administration of Dynamics of Road Transportation Vehicles”. It required road passenger transport enterprises, road dangerous goods transport enterprises, and road freight transport enterprises with more than 50 heavy-duty trucks or tractors to build the vehicle dynamic monitoring platform according to the standards, or use the qualified social satellite positioning system monitoring platform to real-time monitoring and management of the entire road transport vehicles and drivers [2]. However, the construction and operation cost of the dynamic monitoring platform is relatively high, which will inevitably increase the burden on the enterprise. The enterprise will inevitably “cut the work and reduce the material” for the sake of real benefits, weaken the monitoring effect of the dynamic monitoring platform, and lose the platform to improve the safety level of road transportation. In view of this, the concept of vehicle dynamic third-party safety monitoring has been proposed, which is located between the transportation enterprise and the regulatory department. The monitoring is not affected by both parties to carry out safety monitoring work.

Foreign third-party security regulatory research has been applied earlier, and some mature experience have been obtained in some countries. British third-party institutions have originated from Lloyd’s Register of Shipping (LR), which is engaged in inspection, testing and certification services and has been in existence for hundreds of years. Since the 1970s, the British Standards Institute (BSI) has conducted quality management certification [3]. In terms of occupational safety and health management, Australia mainly uses the third-party experts to carry out specialization work. The expert team has the greater authority granted by the local government. In addition to the lack of administrative enforcement power, it can enter the company to conduct occupational health and safety inspections, issue a list of problems, and ask the company to rectify the problems found [4]. For the safety supervision of commercial vehicle, the state of the vehicle is monitored by means of installing a tachograph. Some countries also use the Internet of Vehicles technology to manage passenger coaches. In the management of dangerous goods transport vehicles, for example, Germany requires all dangerous goods transport vehicles must be equipped with a driving recorder and satellite positioning system. The vehicle departure time, travel time and pause time are recorded in the driving recorder and can be recorded for two years [5].

Domestic third-party safety supervision was first applied to coal mines, logistics and other fields [6]. Coal mines are high-risk industries, with many accidents and a large number of accidents. Huang Changsheng et al. (2011) have given the definition of third-party safety supervision: third-party safety production supervision refers to entrusting the safety status in the coal mine production process to the safety management department that is neither the coal mine enterprise nor the coal mine safety supervision institution specially set up by the state. The third-party through the paid way to the coal mine enterprise production safety according to the requirements of the contract to implement supervision and management [7]. The third-party here refers to a safety supervision enterprise established according to the relevant laws of the country, with professional nature, certain authority, independent of coal mining enterprises and government agencies. Third-party safety supervision has also been applied in the construction field. Sizhe studied the game analysis of engineering safety supervision under the participation of third-party safety management. Based on the premise of “incomplete rationality”, modeling and simulating the evolutionary game model consisting of three-party security management, construction unit, and supervision unit [8]. In 2014, Fang of Nanchang University studied the application of third-party agency supervision in food safety supervision, He analyzed the feasibility of establishing third-party institutions to supervise food safety in China and proposed corresponding safeguard measures [9]. In terms of third-party safety supervision of road transportation, the current research is relatively scarce. In terms of application, some regions in China are piloting or formulating programs, and most of them are in their infancy.

This paper analyzes the empirical traffic accident data, as well as the demand investigation of industry management departments and transportation enterprises, combined with the development of vehicle dynamic monitoring technology, proposes vehicle dynamic security monitoring based on third parties, and introduce its pattern implementation framework through the actual case of Xinjiang Uygur Autonomous Region Road Transportation Administrative Bureau. The research results save a large amount of direct costs for transportation enterprises, and at the same time greatly improve the dynamic safety monitoring effect, which has important social and economic significance for the transportation industry of China.

2 Analysis of Problems and Causes of Traditional Vehicle Dynamic Monitoring

2.1 The Main Problems of Traditional Vehicle Dynamic Monitoring

Since the launch of the national key vehicle dynamic supervision work, it has played a big role in curbing traffic accidents, but there are also problems that need to be improved. The following are elaborated from the aspects of industry supervision, enterprises, and Satellite service operators.

Industry Supervision. The problems in the industry supervision are mainly reflected in the shortage of managers. China has a vast territory, numerous road transport enterprises, complex environment and administrative divisions, long transportation

routes, widespread operation across administrative areas and different places. There is insufficient personnel at all levels of traffic safety supervision, and lack of supervision means, resulting in a small number of supervisions not in place. In addition, due to historical reasons, some units have less modern technology applications, and a small number of older people are unskilled in the use of information tools, which also limits the efficiency of supervision.

Enterprises. The first is that the key person in charge of the enterprises have poor awareness. Some may believe that road accident is a small probability event and will not happen in the enterprise. Therefore, they are unwilling to invest and strictly manage in dynamic monitoring. They ignore the emphasis of the management department and the requirements of laws and regulations. The number of full-time dynamic monitoring personnel of the enterprise is insufficient, which makes it impossible for the monitoring platform to be monitored by the 24-hour dedicated personnel or the supervisors are over-fatigued. In terms of personnel quality, the vehicle dynamic related professional knowledge is not well acquired, the requirements of dynamic monitoring of national laws and regulations are not understood, and some even cannot operate the dynamic monitoring platform of the unit. Some enterprise monitors are part-time, and their monitoring time and level are difficult to guarantee.

Secondly, the dynamic monitoring and management system is not perfect and the implementation is not strict. Some enterprises did not follow the prescribed system of vehicle dynamic monitoring system, or did not implement the system.

Finally, there exist problems about monitoring system and platform. The hardware aspects of the monitoring system platform mainly include the monitoring room and computer equipment. Some enterprises do not have an independent monitoring room, and the monitoring room is shared with the front desk, administrative or financial. In terms of hardware configuration, enterprise computers are aging seriously. Some enterprises do not even have special monitoring computers. There are many types of GPS/Beidou vehicle monitoring terminals on the market, the quality is uneven, and some terminals have low good rate, which cannot guarantee real-time online monitoring of vehicles.

The software mainly reflects that the industry management department has basic requirements for the monitoring platform software system, but there are many organizations that provide software systems. Although some of them have the basic functional requirements, the system platform has poor usability and insufficient stability.

Satellite Service Operators. Operators are generally understaffed. Most operators value the market and pre-sales but ignore technology and after-sales, focus on opening up a larger market, and neglect their own technology and service capabilities, resulting in after-sales service cannot support the market.

The internal management and external services of some operators have not formed a sound management system. Marketing, software, hardware and other departments are not well connected, information coordination is not timely, and efficiency is low. There is no unified and standardized service process, and there are fewer channels for customer feedback, and there are no system constraints and assessments for service commitments and complaints.

The hardware of the operator monitoring system platform is mainly poorly configured for the equipment room and server. Some software platform providers are limited to the technical level or service capability and cannot guarantee timely update and upgrade. The system vulnerabilities that occur cannot be solved in time, and do not satisfy the diversified demands of different types of enterprises.

2.2 Analysis of the Causes of Dynamic Monitoring Problems

Comprehensive analysis of the current problems in the dynamic safety supervision of key vehicles, the reasons involved in all aspects of traffic safety management, the following analysis from the industry supervision, road transport enterprises, satellite service operators, and insurance.

Industry Supervision. First, the law enforcement force of the safety supervision team is weak as the insufficiency of employees. Particularly, the staff team is unstable and the law enforcement ability is not strong. Some safety supervisors also undertake other tasks assigned by the unit, and even have multiple jobs, resulting in work fatigue.

Secondly, the means of supervision need to be improved, mainly because the various means, facilities, and equipment required for supervision are outdated. Although the dynamic safety supervision data is backed up to the government's public platform, it lacks corresponding analysis and processing methods, resulting in the data not playing its due role.

Enterprises. Some enterprise management is limited by the concept and cultural level of the business owners. In general, they might be keen on building and approving the plan, but ignore the problems in the dynamic monitoring of the enterprise, solely pursuing economic benefits and ignoring social benefits, or even give up the enterprise related social responsibility[10].

The quality of safety managers and employees are unbalanced. Some safety managers and monitors do not have the knowledge level of safety production management or non-full-time safety manager. It is difficult to achieve a good job in the enterprise safety management.

Some enterprises have backward management concepts, relying only on experience management, only paying attention to punishment but not paying attention to improvement, and treat punishment as management. Even some companies regard some policies as means of revenue, causing practitioners to misunderstand some policies and thus create resistance.

During the period of domestic economic prosperity, the demand for passenger and freight continued to grow. Due to the disorderly expansion and development, the supply of passengers and freight was excessive. In addition, the domestic economic growth slowed down in recent years, and the problem of oversupply became increasingly prominent. Therefore, some road transport enterprises have difficulty in operating or with continued losses, companies are less willing to invest in dynamic monitoring fees or inability to invest.

Satellite Service Operators. Due to fierce market competition, some operators do not hesitate to modify or delete data for enterprises in violation of the rules in order to avoid losing customers.

In addition, there are few operators with strong domestic strength and years of accumulated experience in technology and experience. Some small-scale operators have insufficient experience in road transport industry security monitoring and are at a disadvantage in market competition, which makes their profitability difficult, and thus they are even weaker and inability or no intention to provide better products and services for enterprises.

Insurance. Regulations on Road Transportation of the People's Republic of China clearly stipulates that "Cargo operators and dangerous goods transport operators shall separately insure passengers or dangerous goods for carrier liability insurance." Carrier liability insurance plays a certain role in safeguarding the operational risks of transport enterprises. And played a certain role in the treatment of compensation, resolution of disputes, bus operators and the majority of drivers have a better understanding of the carrier liability insurance, and they have a strong sense of participation.

However, in the fierce market competition, many insurance companies have to pay attention to the performance and compensation while ignoring the risk of prevention and control in order to seize market share. In addition, they are engaged in passive risk management, and more only bear the function of economic compensation afterwards. According to statistics, the cost ratio of many property insurance companies in China is higher than 100%, and the overall profit is not strong. The important reason is that the disaster prevention and loss prevention work is not done well. After the insurance, the risk management is not strengthened and the insurance companies are not carried out daily safety inspections, etc., resulted in high-risk rates and rising operating costs [11]. Compared with some insurance companies with better underwriting benefits in the world, the gap is obvious.

2.3 The Necessity of Vehicle Dynamic Third-Party Security Monitoring

Improve Road Transportation Safety Management and Security System. To achieve the established goals of road transport safety management, we must establish a perfect road transport safety management and security system, and comprehensively consider the five aspects of personnel, vehicles, infrastructures, stations and regulations that affect traffic safety. Only by doing a good job in the safety construction of these five aspects and coordinating the relationship between them and then the road transport safety management work can form a complete guarantee system.

Solve the Problem That Traditional Dynamic Monitoring Can't Solve. According to the analysis, there are still many problems that cannot be solved in the traditional dynamic monitoring. However, after the promotion of the third-party monitoring mode, there are significant changes in the aspects of improving the efficiency of supervision, making up for the lack of safety capacity of enterprises, and optimizing the workflow of supervision.

Take Advantage of Independent Third Parties. The third-party monitoring can improve the road transportation safety supervision system, and can fully utilize the technical and professional advantages of independent third parties, improve the control and incentive mechanism of safety management, promote the transformation of law enforcement concepts, and understand the nature of law enforcement. The new model can save a large amount of originally limited regulatory resources, reduce the personnel and financial burden for enterprises, and conform to the guidelines and policies of the “streamlining administration, delegating more powers to lower-level governments and society, improving regulation and optimizing services” in the new era of the country.

2.4 The Advanced Nature of Third-Party Monitoring

In terms of mode, the third-party monitoring is equivalent to adding a “firewall” between the traditionally monitored enterprises and the monitoring operators, which makes the monitoring data and workflow clear and clear, and eliminates the black box operation, so that the regulatory agency can see it. In terms of technology, third-party monitoring applications use big data and artificial intelligence to analyze all data, realizing the precision and intelligence of security management, and making the data that was originally simply stored have the proper value.

3 Vehicle Dynamic Third-Party Security Monitoring Mode Implementation and Framework

Third-party monitoring of road transport safety production means that road transport enterprises entrust vehicle dynamic monitoring to a third-party safety management department that is neither a safety management department was affiliated to themselves nor a specialized organization of a safety regulatory agency set up by the state. The third-party organization implements dynamic monitoring according to the requirements of the signed production safety contract, focusing on process supervision and real-time warning of violations such as speeding, shielding GPS signals, night driving, no line card operation, and operation beyond the scope of business.

3.1 Mode Implementation and Framework

The third-party security monitoring mode is mainly implemented in the following ways.

One is socialized road transportation safety service agencies or information technology professional companies as third-party security monitoring entities usually require traditional vehicle dynamic monitoring service providers not to be third-party entities. The government formulates the access conditions and management rules for vehicle dynamic third-party security monitoring institutions. The qualified institutions are responsible for the construction of the third-party monitoring platform. Under the coordination of the government, the enterprise directly transmits the original monitoring data of their own monitoring platform or the original monitoring data of the satellite service operator to the third-party monitoring platform through the protocol.

The third-party monitoring platform performs dynamic security monitoring of the enterprise vehicle based on the incoming data. Different monitoring data reports are submitted to the enterprise and industry supervision departments follow the agreed period. The cost of third-party security monitoring platform services is borne by the government through the purchase of services or by the enterprise [12].

The other is the insurance consortium or insurance brokerage company acts as the third-party security monitoring entity. The technical solution of this method is similar to the first one. The main difference is that the monitoring platform is built by the insurance consortium or the insurance brokerage company, the construction and operation costs of the platform are borne by the insurance consortium or the insurance brokerage company. Insurance consortiums or insurance brokerage companies reduce the risk of security accidents by urging enterprises to rectify potential safety hazards, improve their safety level, and reduce accident rates. If the reduced accident compensation costs are greater than the construction and operation expenses of third-party platforms, the insurance associations or insurance brokerage company has the motivation to undertake third-party monitoring.

The third-party monitoring mode can be selected according to the specific conditions in each region. If the two modes are difficult to implement in the region, some adjustments can be made on the basis of the two modes to make it more suitable for local conditions. Figure 1 is a schematic diagram of data transmission between a vehicle dynamic monitoring traditional mode and a third-party monitoring mode.

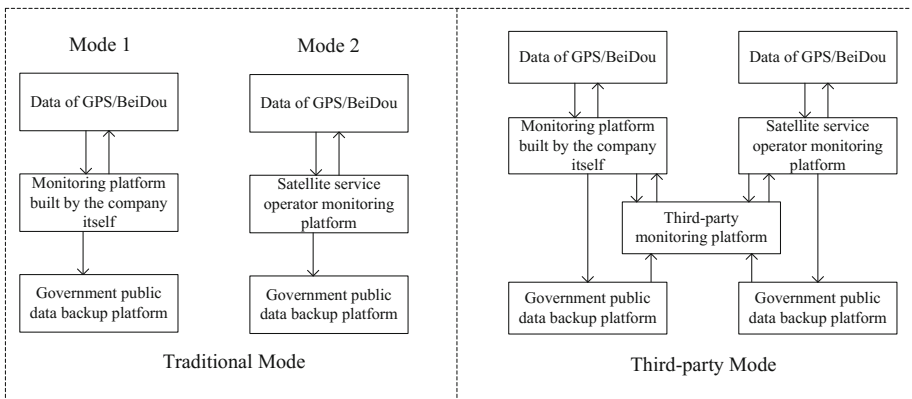


Fig. 1. Schema of data flow between traditional and third-party monitoring mode for vehicle dynamic monitoring.

There are also many differences in the operation mode between the vehicle dynamic monitoring traditional mode and the third-party monitoring mode, mainly reflected in the investment and construction of the monitoring platform, the dispatch of monitoring personnel, the deployment of the monitoring room, and the sharing of expenses, as shown in Table 1.

Table 1. Comparison between traditional monitoring mode and third-party monitoring mode operation

Mode		Platform	Monitoring personnel	Location of monitoring room	Cost
Traditional	Mode 1	Company Self-built	Personnel owned by the enterprise	Enterprise	Enterprises bear the costs
	Mode 2	Rent a satellite service provider platform	Personnel owned by the enterprise or form the satellite service operator	Enterprise or satellite service operator	Enterprises bear the costs
The third-party		Build a completely independent platform	Personnel owned by the third-party	Place owned by the third-party	Government Purchase of Services, Surplus of monitoring costs, Surplus of premium

The surplus monitoring cost refers to the cost of personnel, hardware equipment, energy power and other expenses saved by the third-party monitoring service after the enterprise entrusts the third-party monitoring service; the surplus insurance premium refers to the reduction of the insurance premium of the enterprise due to the improvement of the enterprise’s safety level.

In 2016, Changji Prefecture carried out a pilot project on third-party vehicle dynamics monitoring. After six months of implementation, its third-party monitoring model has gradually matured and achieved certain results. Figure 2 shows the third-party monitoring system platform for vehicle dynamics in Changji.

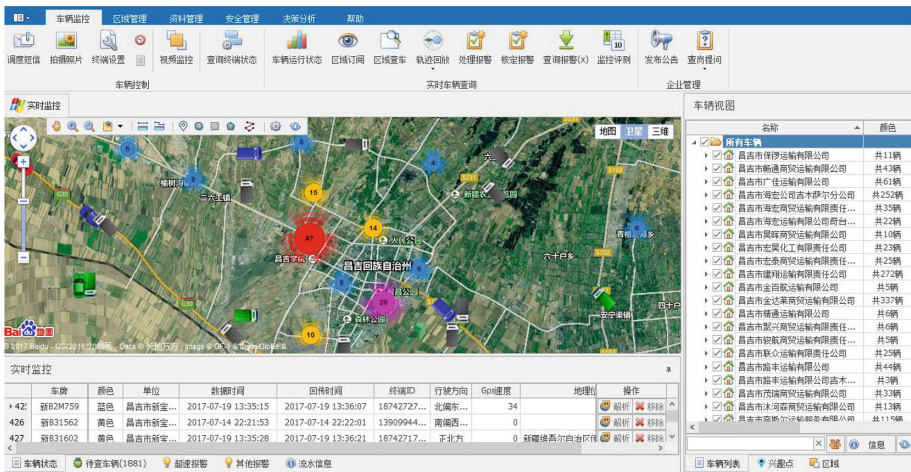


Fig. 2. Third-party monitoring platform interface

3.2 Solved Problems

For Administrative Department. In the past, the dynamic security monitoring work of vehicles was mainly undertaken in enterprises, but the data needs to be filed to the government public platform, and the administrative department could not actively carry out dynamic security monitoring. Allowing enterprises to supervise themselves can basically be done for large-scale and well-managed enterprises. However, for the majority of small and medium-sized enterprises, it is difficult to press the factors such as their safety awareness, personnel investment, and management level. It is stipulated to do dynamic security monitoring. The third-party monitoring organization is based on the supervision of the administrative department, which is equivalent to exercising part of the supervisory duties of the management department. It has no relationship and conflicts of interest with the enterprise, and actively controls the enterprise, plus the pre-risk unique to the third-party monitoring platform. Evaluation and early warning and other functions, the original passive intervention after the intervention as an initiative in the pre-intervention. Third-party security monitoring has greatly improved the regulatory efficiency of the administrative department. It also integrates third-party monitoring data through third-party security monitoring, which lays a solid foundation for vehicle monitoring and unified standardization management and data resource integration application.

For Enterprises. The first is the great promotion of their dynamic monitoring work. Through the standardized dynamic monitoring services provided by third-party monitoring organizations, enterprises can improve their own dynamic monitoring work and learn the advanced experience of third-party organizations, so that the dynamic monitoring level of enterprises can be rapidly improved. Second is the improvement of corporate security. The strict monitoring of third-party organizations has greatly reduced the number of violations of the enterprise, and the awareness of personnel safety has been improved. The facilities and equipment with problems are satisfactory or updated, and the accident rate of the enterprise is gradually reduced. The third is the improvement of the economic efficiency of enterprises. After the implementation of third-party monitoring, relying on the professional services of third-party organizations, enterprises can reduce the input of personnel and equipment in dynamic monitoring. The reduction of accident rate reduces the economic loss of enterprises, and the higher the safety level, the higher insurance premiums rate of discount can be obtained. The preferential rate and high security level can also bring more business and better social reputation to the enterprise.

For Satellite Service Operators. Third-party monitoring does not replace the location of the satellite service operator, and there is no conflict with it, and it will not harm its interests. Driven by third-party monitoring, operators do not have to take the risk of accepting enterprises to revise data and other violation requirements, prompting enterprises to pay satellite service fees on time, through the regulation of the market, greatly reducing unreasonable competition and maintaining the healthy development of operators.

For Insurance. On the one hand, the improvement of the security level of enterprises reduces their insurance premiums. For the insurance industry, the low accident rate is the low loss rate. The road traffic accidents involving casualties usually pay a large amount. Generally, the reduction in premium income is negligible compared to the amount of the payment. Because of this, the insurance industry has the incentive to build and operate third-party security monitoring platforms.

The implementation of third-party safety production monitoring is a way for the “public” (government), “master” (enterprise), and “housekeeper” (socialized third-party regulatory agencies) to win, and is the innovation of the enterprise safety management model.

3.3 Advantages and Case Effects Analysis

As a bridge, the third-party monitoring platform opens up the information channel between enterprises, operators and administrative departments, so that information is no longer one-way transmission, forming a closed loop of transmission-discrimination-feedback. The specialization of third-party organizations is in line with the development direction of professional institutions to do professional work, balances the interests of many parties, and promotes the development of the industry and social progress. Third-party security monitoring implementation does not require large-scale installation of facilities and equipment, nor does it need to replace the original system. It is fully compatible with the original hardware and system. The implementation difficulty is mainly in the early coordination work, and can be quickly implemented under the conditions of agreement. At present, the transfer of some auxiliary functions to third parties (professional intermediaries, industry associations and non-governmental organizations) through business outsourcing has become a major feature of the new public management movements in countries around the world. In various industry categories, more and more third-party supervision models have emerged. Various areas such as food, health, and testing are gradually exploring and promoting third-party supervision models, which are in line with national new policies and international trend orientation.

With the support of the Xinjiang Uygur Autonomous Region Road Transportation Administrative Bureau, the third-party monitoring of road transportation in 2016 began piloting in Changji, and the first time accessed more than 2,000 vehicles from 52 road transport enterprises. The vehicle types covered passenger transport, dangerous goods, general cargo and rental. According to comparative statistics, in December 2015, when the third-party surveillance was not piloted, Changji’s “two passengers and one danger” (i.e., buses, tourist chartered vehicles and dangerous freight vehicles) had more than 280,000 violations of regulations for one month; In May 2017, there were only 2,700 violations of regulations, and the violations of road transport vehicles plummeted, with a proportion of 99% [13]. Before and after the implementation of third-party monitoring, the vehicle-mounted terminals, regulatory standards, and operating environment of the region have not changed. It can be considered that the decline of violations of regulations is almost due to the implementation of third-party monitoring, which excludes the influence of other factors. Figure 3 shows the comparison of the implementation effect data of Changji third-party

monitoring. Third-party monitoring began in September 2016, after which the number of vehicle violations continued to decline, and fell to a minimum in January 2017, and then entered a stable period and the number of alarms remained at a low level.

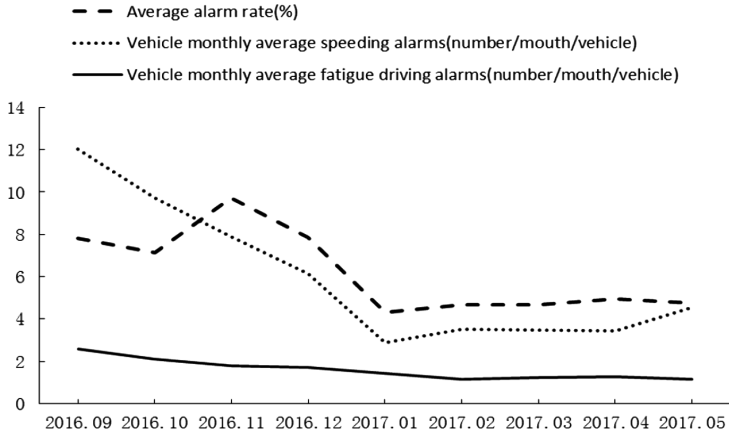


Fig. 3. Comparison of the implementation effect data of Changji third-party monitoring (September, 2016–May, 2017)

According to the statistics of the number of alarms, the number of monthly speeding alarms for bicycles decreased from 12.03 to 4.53, with a decrease ratio of 62.3%; the number of monthly fatigue driving alarms for bicycles decreased from 2.564 to 1.164, with a decrease ratio of 54.6%.

4 Conclusions

The location-based dynamic monitoring of large-scale operating vehicle safety based on GPS/Beidou technology has achieved good results in China, and has curbed the occurrence of serious traffic accidents to a certain extent. It is an important milestone in China’s traffic safety management. The third-party monitoring mode not only continually extends the practicality of dynamic monitoring, but also relies on more professional organizations in technology, so that the monitoring platform, staffing, and hardware facilities are always in an advanced state. Management is also based on the role of a third-party independent identity, so as to avoid interference from the regulated party, in order to conduct objective and fair monitoring and management. Based on previous experience, some areas are currently conducting third-party security monitoring for road transport deployment. It is believed that with the full implementation of third-party security monitoring, China’s road transport safety level will be further improved.

Through the research on the demand of the transportation industry administrative departments and enterprises, combined with the development of the existing vehicle

dynamic monitoring system technology, this paper proposes the dynamic security monitoring of vehicles based on third parties, and opens up the information channels between enterprises, operators and administrative departments. The information is no longer a one-way transmission, but a closed loop of transmission-discrimination-feedback is formed, and its implementation framework is introduced through the actual case of the Xinjiang Uygur Autonomous Region Road Transportation Administrative Bureau. Comparing the violations of vehicles before and after the implementation of third-party monitoring, it was found that the violations of road transport vehicles decreased significantly, with a proportion of 99%. The research results of this paper save a lot of direct costs for transportation enterprises, and also greatly improve the dynamic security monitoring effect, which has important social and economic significance for China's transportation industry.

From the perspective of future development trends, third-party monitoring platforms need to integrate more monitoring information and gradually reach the coverage of holographic monitoring. Based on the Internet of Vehicles technology, vehicle information can be transmitted to the monitoring platform for better analysis of vehicle status. Combined with the vehicle video terminal, the advanced image analysis technology is used to identify the driver's behavior status in real time, prevent and warn the driver's unsafe behavior and assist the driver in driving and emergency. Through information fusion and big data technology, it is possible to accurately grasp and accurately predict the state of the vehicle and the driver, and improve the safety level from the essence [14, 15]. Of course, the development of unmanned vehicles in the future may fundamentally solve many of the current transportation safety problems.

Acknowledgments. The authors are thankful to the Shanxi Bei'an Fire Protection Technology Co., Ltd. for providing the necessary data and information in this study. The research was sponsored in part by the Shanxi Key R&D Program (High-Tech) Projects (No. RCS2018K010), Science & Technology Department, Shanxi Province, the Humanities, and Social Science Research Project (No. 18YJCZH011), Ministry of Education, China, and the Shanghai Municipal Natural Science Foundation (No. 17ZR1445500). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

References

1. National Bureau of Statistics, Statistical communique of the People's Republic of China on national economic and social development in 2016 (EB/OL). http://www.stats.gov.cn/tjsj/zxfb/201702/t20170228_1467424.html. Accessed 28 Feb 2017
2. Order of Ministry of Transport of the People's Republic of China, The Ministry of Public Security of the People's Republic of China, State Administration of Work Safety no. 5 of 2014, Measures for the Supervision and Administration of Dynamics of Road Transportation Vehicles (EB/OL). Accessed 28 Jan 2014
3. Li, H.: The service safety regulation of third-party institutions in the UK can be used for reference. *Labor Prot.* **5**, 96–97 (2016). (In Chinese)
4. He, Z.: An overview of occupational safety and health social services abroad. *Chin. Occup. Saf. Health* **3**, 60–61 (2016). (In Chinese)

5. Tang, L.: Study on the Key Technologies of Safe Supervision for Commercial Vehicle Operations Supported by Information Technology. Chongqing University, 2012 (In Chinese)
6. Wu, J., Peng, Q.-y.: Research on the third part logistics project management maturity model. *J. Transp. Eng. Inf.* **4**(3), 87–92 (2006). (In Chinese)
7. Huang, C., Ma, H., Chen, C.: Mechanism about introducing third party supervision in coalmine safe production. *Henan Sci.* **29**(7), 878–882 (2011). (In Chinese)
8. Pan, S.: Game Analysis of Construction Safety Regulation with the Participation of Third Party's Safety Managements. Huazhong University of Science and Technology (2015). (In Chinese)
9. Fang, Y.: Third party supervision agencies-study on China's food safety supervision model. Nanchang University (2014). (In Chinese)
10. Sun, D.J., Zhao, Y., Lu, Q.: Vulnerability analysis of urban rail transit networks: a case study of Shanghai, China. *Sustainability* **7**(6), 6919–6936 (2015)
11. Sun, D.J., Guan, S.: Measuring vulnerability of urban metro network from line operation perspective. *Transp. Res. Part A: Policy Pract.* **94**, 348–359 (2016)
12. Yang, L., He, B.: An analysis on the influence factors of technical progress rate for insurance companies in China: Based on malmquist index and random-effect panel data model. *J. Zhejiang Univ. (Humanit. Soc. Sci.)* **7**, 1–17 (2016). (In Chinese)
13. Transport Department Of Xinjiang Uygur Autonomous Region (EB/OL). <http://www.xjjt.gov.cn/article/2017-6-13/art132306.html>. Accessed 15 July 2017
14. Huang, Y., Sun, D.J., Zhang, L.: Effects of congestion on drivers' speed choice: assessing the mediating role of state aggressiveness based on taxi floating car data. *Accid. Anal. Prev.* **117**, 318–327 (2018)
15. Zhou, J.: Discussion on the third-party safety supervision mode of interprovincial passenger transportation. *Commun. Shipp.* **4**(2), 55–58 (2017). (In Chinese)



Research on Application of BIM Technology in Municipal Engineering Construction

Xiaoqing Zeng¹, Qipeng Xiong^{1,2}, Yizeng Wang^{3(✉)}, Xinchun Xu¹, and Liqun Liu¹

¹ The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, No. 4800 Cao'an Road, Shanghai 201804, China

² Shanghai FR Traffic Technology Limited Corporation, Shanghai, China

³ Shanghai University, Shanghai, China

1102790744@qq.com

Abstract. At present, the underground space tunnel in municipal engineering is the advanced layout of technical facilities and pipe network. And it is also the trend of urban construction and development. The main difficulty of underground space tunnel construction is the limited space and complicated pipelines in cities. In the past, during the construction process of metros, utility tunnels and ground crossing signal control, there were pipeline collisions and thus on-site changes are required. Those changes often make local space localized and unreasonable, because of the collision of large-section ventilation ducts. In order to effectively respond to the requirements on promoting the application of BIM (Building Information Modeling) from the Ministry of Housing and Urban-Rural Development, this paper studies the establishment of the visual model of engineering and surrounding environment, the data acquisition mechanism for metro BIM systems, and the collaborative information management platform for metro design based on BIM. The research results have the effect of shortening the construction period, controlling the cost, and ensuring the quality of the design of the utility tunnel. The research results can provide guiding significance for BIM applications on metro design in the future.

Keywords: Metros · Utility tunnel · Signal control at intersection · BIM technology · Data visualization · Collaborative information management

1 Introduction to BIM Technology

BIM (Building Information Modeling) is a data tool applied to engineering design and construction management. It integrates various project information through parameter model, then share and deliver it in the whole life cycle of project planning, operation and maintenance. So technicians can correctly understand and respond to various building information. It also provides the basis for collaborative work for the design team and all parties involved in the construction operation unit in order to improve production efficiency and shorten the construction period.

There are numerous pipelines and support hangers inside the metro and utility tunnel. The intersection and the branch of the utility tunnel have relatively complicated pipeline connections and civil structures. It is difficult to be accurately expressed by the traditional two-dimensional design. The BIM technology can directly give the three-dimensional information model of the metro and the utility tunnel. It provides more intuitive design results and clearer design intent.

BIM technology makes people know what the metro or the utility tunnel looks like after its completion during the design stage. Based on it, the project can be demonstrated and adjusted. Problems can be found and solved in time. Through the BIM modeling of each major, the design and construction plan will be simulated and previewed in the BIM system, so as to reduce the reworking and shorten the construction period.

In addition, because the BIM model contains a wealth of information, it can provide data for professional analysis and simulation, so as to draw accurate conclusions and make the project demonstration and construction explanation more efficient. At the same time, the large amount of data generated during the design process will play an irreplaceable role in the later construction, operation and even the whole life cycle of engineering management.

2 Visual Modeling

2.1 BIM Establishment of 3D Information Model for Metros and Utility Tunnels

It is easier to express the design concept and understand the design content by 3D drawing of Metros and utility tunnels. We can analyze the tunnels' internal civil structure and various pipeline correlations through 3D model.

BIM can directly show three-dimensional information model, provide design results intuitively, and describe complex problems clearly. So those non-professionals can also clearly understand the design with no need for imaging what the two-dimensional drawings look like after the completion of the project.

2.2 Simulation of Metro and Utility Tunnel Design Scheme

Through BIM visualization technology, the design concept and design scheme of Metro and utility tunnel are displayed visually. And it is easy to choose the optimal design scheme which has advantages of reasonable structure, economic benefit and social benefit by scheme demonstration and scheme comparison.

2.3 Collision and Clear Height Check of Engineering System and Surrounding Environment

Visual model is established to simulate the construction on the virtual platform to check the collision between pipelines, hoisting and surrounding buildings. It is available to check the civil structure, pipeline installation and working place in the utility tunnel to

optimize the spatial layout. Visual model is also can find the problems of utility tunnel standards, feeding inlet, air outlet and the clearance height of all kinds of pipelines during design process in order to avoid reworking.

2.4 Reserved Hole Inspection Based on 3D Information Model

There are many kinds of pipelines in metros and utility tunnels, and there are many intersecting nodes between pipeline and civil structure. It is often happened that reserved holes is not buried in civil structure. While 3D information model can be used for checking reserved holes automatically to avoid similar mistakes.

2.5 Virtual Construction of Some Complex Nodes

Complex nodes are simulated in BIM platform. Then construction schemes are compared to get familiar with construction steps. Thus low-level errors and blindness in construction process can be reduced and curbed.

2.6 Accurate Calculation of Engineering Quantity

In traditional design mode, the engineering quantity is mostly estimated by cross-section method. While the error is always great when the topographic relief is large or the plane is complex. And the workload of pipeline engineering is mostly estimated according to standard drawings and experience, therefore the error is relatively large.

But in BIM design model, because of accurate geometric and attribute data in each sub-model, it can get accurate volume, area, length, Engineering quantities and even the accurate engineering quantities of support model with the help of professional calculation tools.

3 Data Format

3.1 Theoretical Research on Data Acquisition Mode

Data are from survey and design documents and various monitoring data on site.

3.2 Theoretical Research on General Data Standard

The data format, logical relationship and expression form of all kinds of data in the early stage of consulting design are studied to form a data standard which is extensible and independent of the platform. It means that almost all operating platform can be based on this database.

4 Life Cycle Management Platform

4.1 Platform Function at Design Stage

(1) Design assistance and coordination plan

Two-dimensional plane, vertical or sectional drawings can be quickly gained by using the three-dimensional model to ensure the accuracy of the drawings. Thus it can significantly reduce the workload of designers and proofreading.

On the premise of complete professional drawings, the project planning of different installation units and different work interfaces can be merged into the same model. Then it can show the sequence of construction and installation intuitively. And it is convenient to check the rationality and practicability of the plan which contains different working-procedures and construction. It can also carry out construction planning coordination and arrangement more accurately to reduce time wasting. Through system synthesis, the system design is optimized, and the problems in the design are found in advance. For the nodes with relatively complicated working conditions, the three-dimensional node diagram can be generated in real time, and the coordinated installation measures can be ensured in place.

(2) Engineering quantity statistics

The project participants, such as joint-design, bidding agency, investment supervision and so on, are led by the proprietor to check the engineering quantity and control the cost before bidding.

4.2 Platform Function at Construction Stage

(1) Deepening the design conflict detection scheme

Models are modeled on real scales, and some parts omitted in traditional expressions (e.g. pipeline insulating layer) are displayed. Thus it can reveal some deep-seated problems which are seemingly without problems (Figs. 1 and 2).

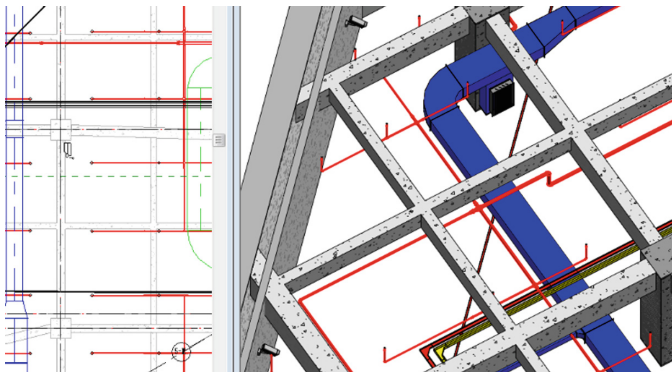


Fig. 1. Plane drawings correspond to 3D models

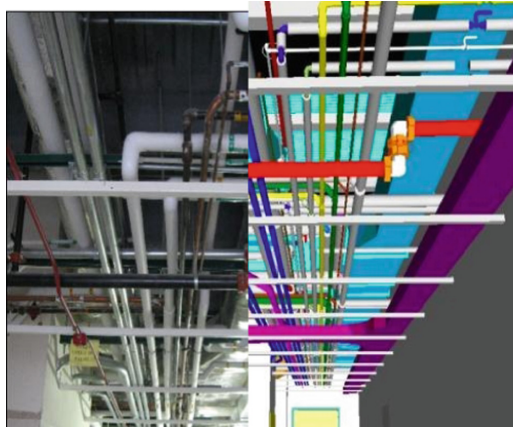


Fig. 2. 3D model and the real project comparison

(a) Collision detection and rational layout

The traditional two-dimensional drawings cannot fully reflect the possibility of collision between individual and professional systems. At the same time, because the discrete type of two-dimensional design is unpredictable, designers will also miss some problems of pipeline collision. The application of BIM technology can make full use of its collision detection function to feed back the collision point to the designer as soon as possible. Communicating with the proprietor and consultant timely can minimize the collision and rework of the pipeline at the deepening design stage.

(b) Equipment parameter re-check and calculation

In the traditional deepening design process, the system parameters are checked and calculated based on the two-dimensional plane picture. That is almost different from the actual installed system, which can result in inaccurate results. Now using the BIM, we just need to click a mouse after drawing a good mechanical and electrical system model. It can let the software automatically complete complex calculation work. If the model changes, the result will also be updated.

(2) Monthly progress quantity re-check plan

After the establishment of the model, according to the relevant rules of the contract, the quantities of each subsystem are collected. Then the detailed quantities of the section are given, and the monthly quantities schedule is calculated according to the monthly project planning (Figs. 3 and 4).

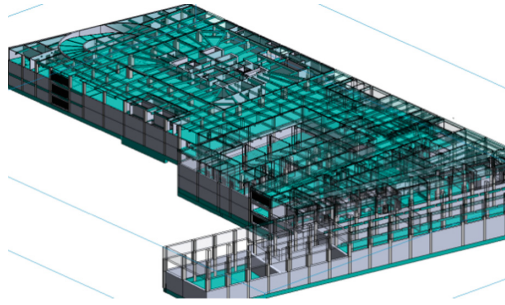


Fig. 3. Project quantity plan at the beginning of the month

At the monthly work summary stage, the actual monthly completion of the project schedule is reviewed by BIM model. So, the accurate monthly progress of the project is obtained.

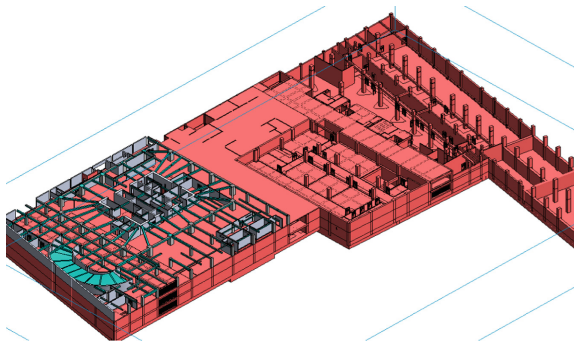


Fig. 4. Actual completed quantities at the end of the month

At the settlement stage, the engineering quantity can be judged and verified intuitively according to the actual construction technology, construction process and construction boundary conditions.

For the changes in the construction process, it can also count the detailed engineering quantities at the first time.

In addition, according to the construction drawings, drawings changes and so on at the project planning stage, it can reflect the amount changes of relevant projects. BIM can provide dynamic construction drawings budget with detailed project planning, which is easy for dynamic investment control.

(3) Simulation plan of construction progress and complex process

4D virtual construction based on BIM adds time into the 3D model completed in the design phase to form a 4D simulation animation. We can timely discover the possible risks, and adjust the model and plan according to the problems. Even if there

are some design changes, construction drawings changes, etc., schedule modification can also be quickly changed. In addition, the 3D model and virtual animation can make the evaluation experts know the arrangement of construction organization, main construction methods and overall plan of the project. So the experts can make a preliminary evaluation of the construction experience and strength of the bidder.

(4) Construction site cooperation plan

After the establishment, examination and virtual acceptance of the model, we have a general model which can be used to guide the construction. Then according to the detailed project planning, we add corresponding time attributes to the main components of the model, so that the planning and the model are closely combined. The construction sequence is now being planned and presented by the most realistic 4D model.

(a) Quality and safety management

The cloud platform can carry out the real-time transmission of Engineering information, such as text, voice, pictures, personnel signing-in, information network publishing, efficient emergency command and other functions. It is convenient for managers to use fragmented time to grasp the first-hand engineering information in real time. Daily work can be recorded for tracking.

(b) Monitoring data management

On the basis of the system, through the secondary development, the construction monitoring data port is opened. All kinds of automatic monitoring data are introduced, and three-dimensional display is carried out through the platform.

In the platform, the parameters can be recorded in the corresponding components. And the data can be saved and transmitted.

Through the data in the model, the time history curve can be drawn with automatic alarm. It can also realized visual automatic monitoring and safety warning in construction process.

(c) Integration of personnel management and systems

On the one hand, the system can be associated with the site entrance guard management system to display the access card data, which is displayed in forms of lists and curves.

On the other hand, we can use the cameras in construction site to achieve real-time monitoring (Figs. 5 and 6).



Fig. 5. Amplification of video surveillance window associated with BIM model

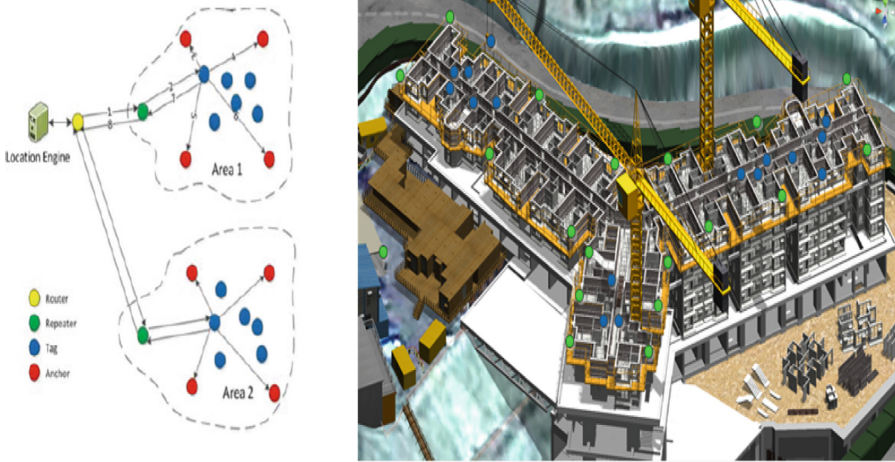


Fig. 6. The combination of safety helmet positioning principle, personnel positioning and BIM model

In addition, if the safety helmet monitoring system is introduced into the project, the combination of BIM technology and safety helmet positioning system can be studied to ensure the sufficient safety of constructors.

(5) Completion acceptance plan

(a) BIM application workflow at completion stage

See Fig. 7.

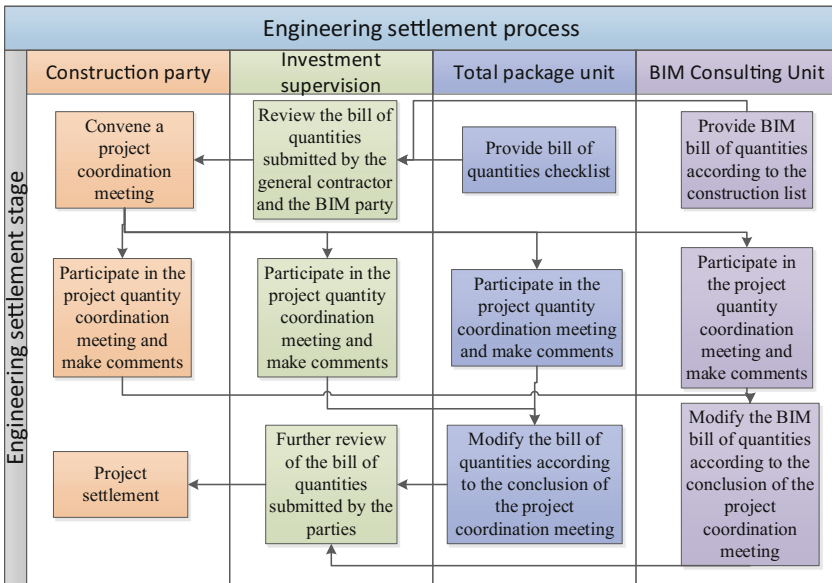


Fig. 7. Flow chart of engineering quantity settlement

(b) Model file designation

The naming of the model file is as follows: prefix-retain-suffix (annotation)-extension

(c) Model delivery standard**① Scope of application**

The model delivery is divided into three stages at the stage of construction and operation.

- model delivery in the feasibility study report phase
- model delivery at pre-bidding stage
- model delivery at construction stage

② Delivery of BIM model at the stage of feasibility study report includes scheme model, ground environment model, underground environment model, model deliverer, model Receiver and model delivery form.

③ Delivery of BIM model at pre-bidding stage

- bidding model: divide the model according to the bidding segment, and finally integrate the models
- environment model
- list of quantities calculated according to the model
- model delivery form

④ Delivery of BIM model at construction stage includes implementation model, environment model, list of quantities and model delivery methods.

4.3 Platform Function at Operation and Maintenance Stage**(1) Integration and transfer of data**

When the constructor transfers the construction products to the maintenance management, that is completion stage, BIM project management team provides not only conventional completion drawings but also an integrated BIM completion model. The project participants, such as joint-design, construction side and so on, are led by the proprietor to establish the completion information model according to the completion drawings and the data in the construction process. Finally the project completion materials are digitized.

After the model is filed, through the keyword fuzzy search, it can quickly find a certain kind of equipment containing 'keyword' in the whole system. Clicking the equipment button in the form can realize the equipment search positioning and enlargement to query the equipment construction information, product description, operation manual and so on.

At the same time, the asset coding information in the model is handed over to the asset manage side, so that the list of assets is clear. The related assets and equipment can be queried and located in real time, so as to realize the information and visual management of all assets.

(2) Establishment of an intelligent operation and maintenance platform

Cooperating with operation and maintenance side, it establishes a BIM-based operation and maintenance management platform. Then it integrates all kinds of functions according to the overall operation and maintenance of equipment, space, energy and so on. Basic technical requirements must be satisfied as followings:

- to query the whole project information of different buildings by 3D GIS system
- to inquire the municipal pipelines of the whole project and the quick connection with other places by GIS system
- to take the model as an auxiliary data for operation and maintenance of the interior space in buildings
- to manage the use of the interior space in buildings by the model
- to show the energy consumption of different parts of the building through the model
- to assist equipment operation and maintenance by the model

(3) Development of intelligent parking management system

The parking management system based on GIS is developed by cooperating with the operation and maintenance company (Fig. 8).

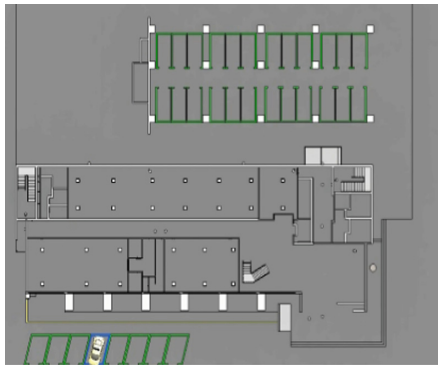


Fig. 8. The intelligent parking management system

(4) Optimization of emergency plans

Cooperating with operation and maintenance management team, it can simulate the emergency plan and verify the operability, rationality and effectiveness by BIM model to provide effective suggestions for optimization and improvement of the emergency plan.

(5) Optimization and improvement plan for operation and maintenance

On the basis of understanding the relevant process and requirements of the project maintenance, the BIM method is embedded in the previous work process to assist its maintenance. It makes maintenance more convenient and orderly. It mainly involves three aspects as followings.

- **Information interaction between device maintenance and BIM 3D model**

Maintenance plan is put into the corresponding model components of the equipment, and maintenance plan can be automatically generated. Maintenance records can be uploaded to the information model by the mobile terminal. It can generate a medical record card to achieve information exchange of information model and daily maintenance work. And the maintenance Electronic Records (text, pictures, sound) can generate electronic maintenance files for the whole life cycle of the metro.

- **Automatic monitoring of underground structure safety**

Introducing the automatic monitoring technology into the operation and maintenance monitoring of the underground structure, it can show the deformation and settlement of the underground structure in real time.

At the same time, it can directly reflect the relevant monitoring data and highlight the parts of large changes by linking the automatic monitoring data into the model. Then it can make early warning or alarm, which can greatly reduce the risk of tunnel operation.

- **Visual display of facility operation**

It can link the monitoring data of existing facilities when the existing facilities system can provide the corresponding interface and data reading in order to realize the visual display of equipment early warning and rapid positioning query (Fig. 9).

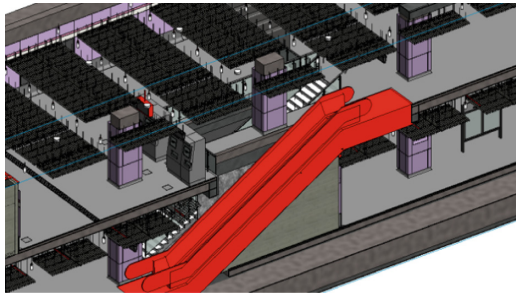


Fig. 9. Elevator failure alarm

5 Conclusions

In this paper, BIM technology is used in the design phase to analyze and study the visualization model of the project. The research focus on metros and utility tunnels to study its surrounding environment, the data acquisition mechanism of BIM system and the design collaborative management information platform based on BIM. Considering all kinds of problems that may exist in the later period, the simulation and solutions are carried out at the beginning to optimize the design depth, so as to shorten the construction period, control the cost and ensure the quality of the project. It is of great

theoretical significance to control the design quality, time limit and the cost in advance in all aspects during the design process. The research results will be of guiding significance for the future application of BIM at the design stage of metros and utility tunnels.

References

1. Wang, G., Zeng, X., Bian, D., et al.: Research on modern tram auxiliary safety protection technology based on obstacles detection. In: Proceedings of the International Symposium for Intelligent Transportation and Smart City (2017). Springer. EI:20171603584789
2. Zeng, X., Wu, C., Chen, Y., et al.: Research on the model of traffic signal control and signal coordinated control. In: Proceedings of the International Symposium for Intelligent Transportation and Smart City (2017). Springer. EI:20171603584791
3. Wang, G., Zeng, X., Yuan, T. (eds.) Study on the influence of train control system on service quality of rail transit. In: International Conference on Service Systems and Service Management (2017)
4. Jian, L., Xiao-Qing, Z., Tuo, S., et al.: Design of the fault recording function in a railway signal microcomputer monitoring system. In: Proceedings of the CICTP 2016 Green and Multimodal Transportation and Logistics 16th COTA International Conference of Transportation Professionals, 6–9 July 2016, Reston. ASCE - American Society of Civil Engineers (2016)
5. Fang, Y., Zeng, X., Zhang, C.: Safety assessment approach for onboard ATP system of Changsha low-speed Maglev project. In: The 23rd International Conference on Maglev 2016, Beilin (2016). 9783940685285



Study on BIM Technology Application in the Whole Life Cycle of the Utility Tunnel

Chuanpeng Hu^(✉) and Shilang Zhang

China Energy Engineering Group
Guangdong Electric Power Design Institute Co., Ltd., Guangzhou 510663, China
331902172@qq.com

Abstract. Building information model (BIM) technology has made considerable development in recent years and has a tremendous impact on the development of engineering construction field. In order to improve the efficiency of design and construction of the conventional utility tunnel, by studying the problems existing in the traditional design process and construction management mode of the utility tunnel, the design and construction integrated optimization model based on BIM is established. On the other hand, in order to solve the problems such as low efficiency and management difficulties existing in the traditional maintenance mode of the utility tunnel, a smart operation and management mode based on BIM technology was established through the study on the application of BIM technology to the operation and maintenance management system of the utility tunnel. Combining the integrated optimization mode of design and construction with the smart maintenance mode, the visualization, integration, scientific and intelligent management in the whole life cycle of urban utility tunnel can realize in some degree.

Keywords: Utility tunnel · BIM · Maintenance · Whole life cycle · Integrated

With the rapid development of cities and the high concentration of population, urban underground space development is more and more important, and the utility tunnel, as the “blood vessel” of the underground space, also receives more and more attention. For a long time, due to the backwardness of technology and obsolete management philosophy, the design and construction of the utility tunnel have been in a state of isolation from each other and the effective information communication way cannot be established between them. On the other hand, due to the short of scientific management and reasonable arrangement, a lot of utility tunnels are in a state of extensive operations after completion. For the existing problems of traditional utility tunnels in the design, construction and operation, it is especially necessary to find a reasonable and effective way to get through the whole utility tunnel life cycle from design to operation management. The rapid development of internet technologies, such as big data technology and of artificial intelligence, has brought new changes to the traditional construction industry, and building information model (BIM) is one of them. The appearance and development of BIM technology has a significant influence on the development of engineering construction field [1]. The core concept of BIM is to focus on the information transmission, sharing and feedback throughout the whole utility tunnel life

cycle from design to maintenance to realize collaborative operation. Therefore, it is significant to apply the BIM technology in the construction of urban utility tunnels, in order to make the design, construction and operations of the utility tunnel more collaborative, and to improve the efficiency of the utility tunnel construction.

1 Study on the Traditional Collaboration Mode of Utility Tunnel Design and Construction

Utility tunnel, also called the municipal pipeline, includes engineering structures and attached facilities which are built in urban underground to hold two or more urban pipelines [2]. Utility tunnel can make all kinds of municipal pipelines, such as water supply pipe, sewage pipe, gas pipe, and electricity pipe, intensively installed, which forms the intensive construction and management and achieves the comprehensive utilization of underground space and the sharing of resources. The utility tunnel construction involves lots of cross-disciplinary efforts, therefore efficient and effective communication between different professions is an important premise to ensure the quality of the whole utility tunnel engineering.

1.1 Analysis of the Collaboration Problem Between Different Professions in the Design Phase of the Traditional Utility Tunnel

According to the difference of professions, the design of utility tunnel can be divided into architecture design, structure design, and MEP (mechanical, electrical, and plumbing) system design. Based on the research of the traditional utility tunnel design process, it can be found that different professions are completely independent, no matter in the schematic design phase, the preliminary design phase, or in the construction design phase and the whole design process and information transfer basically present a single linear status. The adjustment information of upstream profession can not be passed to the downstream profession in time, which causes the design confliction between different professions can not be found until the final summary phase, resulting in huge rework effort. To sum up, utility tunnel engineering problems existing in the traditional design pattern are as follows:

- Information transfer

There is almost no horizontal information transmission between different professions, and nearly all materials are existing with CAD drawings and text data, which can not fully reflect the state of design data and can easily lead to information loss.

- Specialty cooperation

The design mode of each specialty is stale and there is very little correlation between each other. The split between architecture, structure, and MEP is getting more and more serious, leading to large amount of rework and waste of human resources.

1.2 Analysis of the Collaboration Problem Between Different Units Involved in the Construction Phase of the Traditional Utility Tunnel

In the traditional construction mode, the construction process is working in stages and the information transmission in each stage takes an approximately linear form. Technical interpretation of design intention, dominated by design units, is usually based on two-dimensional CAD drawings, and the construction plan has to be formulated in a very short time. During the implementation phase of construction, on-site construction management department needs to coordinate the various type of workers, to control quality, progress, and safety problems of the construction site, to timely solve the construction material supply problem. To sum up, the traditional utility tunnel construction mode has the following three shortages:

- Construction information carrier

Information communication during the whole construction process of the utility tunnel is based on two-dimensional CAD drawings, which brings certain difficulties to the owner units and construction units. As a result, some design problems are hard to be discovered, and the security risk of the construction implementation phase is increasing.

- Construction information transfer

In the traditional construction mode, information is linearly delivered during the construction preparation phase, implementation phase and the phase of final acceptance. There is no the information feedback mechanism, which directly leads to the lag of problems discovering and solving.

- Construction participants collaboration

The participation of design unit, supervision unit and owner units during the utility tunnel construction process is not enough, and there is no effective information communication channel established between each other. Furthermore, inside the construction unit, different types of labor is also mechanically connected in series, with the lack of cooperation.

1.3 Analysis of the Link Problem Between Design and Construction of the Traditional Utility Tunnel

Due to professional division, the design and construction of the traditional utility tunnel project is during a long-term separation, and the only information carrier of the cohesion process is the two-dimensional CAD drawings. In this mode, the design unit takes the lead and design information only unidirectionally flows from design unit to construction unit in the process, which design leads to the defects of design phase building up to the construction phase. The construction problem caused by design defects not only affects the construction process, but also virtually increases the utility tunnel engineering construction cost. Thus, it is obvious that the link between design and construction of the traditional utility tunnel has the following deficiencies:

- The constructability of the design

During the one-way transmission of information flow from design to construction, sometimes designers are lack of practical engineering experience, which causes the constructability of the design decrease. These problems can only be found in the process of construction;

- Separation of design and construction

Design units and construction units just rely on a two-dimensional construction drawings for information transmission, which cannot effectively passes complete design information to the construction unit.

2 Study on the Collaboration Mode Based on BIM of Utility Tunnel Design and Construction

2.1 The Collaboration Mode Between Different Professions Based on BIM in the Design Phase of the Utility Tunnel

The collaborative design of engineering and construction refers to a kind of collaboration mode between multi-specialty in which all the professions inside the design unit and related design management professionals are working in the same design platform and sharing the same real time information data, in order to solve the professional design conflict problem caused by the lack of relevance [2].

The collaboration problem between different professions in the design phase of the traditional utility tunnel mentioned in Sect. 1.1 can be solved by using BIM technology. The design progress information of each professional can be synchronized to the BIM center file of the utility tunnel. From this BIM center file each profession can obtain the relevant data they require, which breaks the original design pattern and sets up a new collaborative design model based on BIM technology of utility tunnel as shown in Fig. 1.

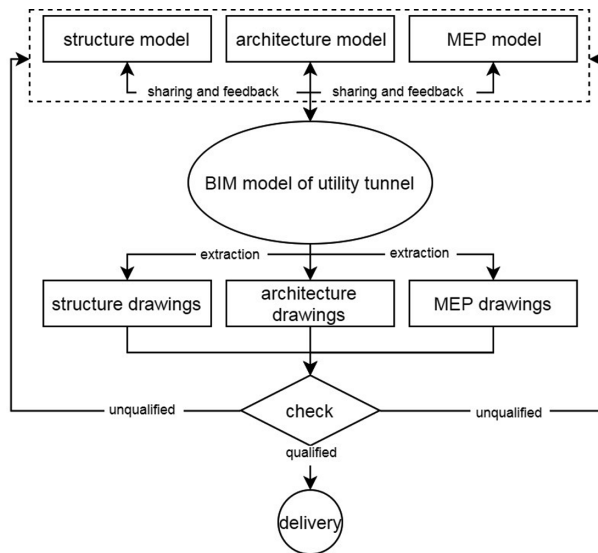


Fig. 1. The collaborative design model based on BIM technology of utility tunnel

Compared with the conventional collaborative design mode, all professions start their work almost at the same time in this new mode based on BIM technology. The structure profession and MEP profession begin their work much earlier than before. Due to the BIM center file and linked mechanism, the horizontal flow of the design information can be realized. On the other hand, design management department can ensure the vertical flow of the design information by inspecting the BIM center file directly. The horizontal and vertical information flow model reduces not only the design processes, but also the frequency of information transfer, and makes the cooperation between different professions working throughout the entire design process. As a result, the design problems can be discovered in time, feedback timely and solved in time, greatly improving the design of cooperative efficiency.

2.2 The Collaboration Mode Based on BIM Between Different Units Involved in the Construction Phase of the Traditional Utility Tunnel

Coordinate construction refers to a kind of collaboration construction mode in which the construction management department acquires the real-time data and information and processes these data in time during the whole process of construction project construction, and efficient information exchange channels between different professions can be built up, which breaks the “information island” phenomenon and improves the efficiency of construction management synergy [3–5]. For the collaboration problem between different units involved in the construction phase of the traditional utility tunnel mentioned in Sect. 1.2, the BIM model of the utility tunnel can be used to integrate different construction units together and establish the information transfer and cooperation mode with a core of the BIM model as shown in Fig. 2.

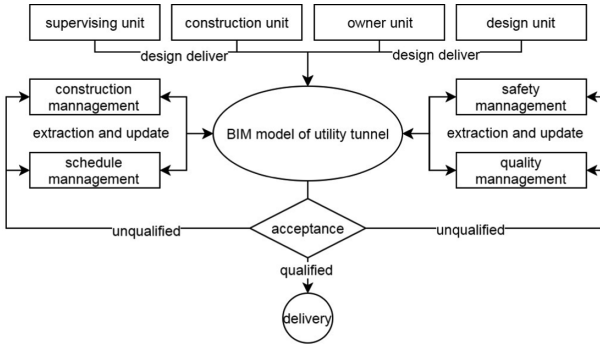


Fig. 2. The collaborative construction model based on BIM technology of utility tunnel

Compared with the conventional collaborative construction mode, the BIM utility tunnel model is playing a vital role throughout the whole process including technical interpretation of design intention, implementation and the final acceptance. During the whole construction process, all the units involved exchange information and collaborate with each other based on the BIM model. By weakening and blurring the boundary of different construction phases, the BIM model of the utility tunnel can make the whole process including technical interpretation of design intention, implementation and the final acceptance form an organic integration. In the process of utility tunnel construction, construction management department can use BIM technology to track the real-time progress of construction, to predict the possible problems the construction units may face and to prepare in advance. By using the sharing information of the BIM model, the different labors inside the owner units, design units and construction units can be connected with each other effectively so that all the contractors of the utility tunnel project may have a better understanding about the project and cooperate with each other efficiently. Furthermore, the key information can be shared appropriately, which ensures the technology and information support for the implementation of the project.

2.3 Design and Construction Integrated Optimization Model Based on BIM

In order to solve the problems caused by the separation of the traditional design and construction, this paper brings BIM technology into the whole process of the utility tunnel engineering. Using the BIM model of the utility tunnel as a linking bridge between design and construction, the design and construction integrated optimization model based on BIM can be established. Compared with the conventional cooperation model between design and construction, the optimization model based on BIM can realize the two-way transmission between design units and construction units. Therefore, with the aid of the three-dimensional view of the BIM model, construction units can understand the design concept more clearly and achieve a seamless link with design units.

The design and construction integrated optimization model based on BIM is realized by using BIM center file as a medium, which achieves seamlessly link from design to construction. The design results no longer flow unidirectionally and irreversibly from design units to construction units which blurs the boundary between design and construction and realizes the deep fusion of the two phases. Therefore, due to the construction contractors involved in the design phase in advance, they can have a deeply understanding of design concept, which also improves the constructability of the design [6].

3 Study on the Smart Maintenance Mode of Utility Tunnel Based on BIM

From the analysis above, it is easy to get the conclusion that the conventional design, construction and maintenance mode are unable to ensure the effective and convenient management of the utility tunnel during the whole design working life about 100 years. Using the BIM model of the utility tunnel as the core, the whole information can flow unobstructed between different departments. Every manager can get what they really need from the model conveniently and they in turn can also update the latest information to the BIM model. This iterative mechanism plays a vital role in the healthy operation of whole BIM system [7].

3.1 Analysis of the Problems Existing in the Maintenance of the Traditional Utility Tunnel

For the conventional maintenance of the utility tunnel, the completion drawing of project always be handed over incompletely and the accurate information of utility tunnel structure and associated equipment cannot be obtained. Since all drawings and data are lack of relevance, the accuracy of facilities management is affected inevitably [8]. If the data about decoration, improvement, or inspection is unable to be updated in time, the effectiveness of the management system will also be influenced. Furthermore, the unintuitive CAD drawings and insufficient data feedback may lead to the distortion of the data and analysis results [9, 10].

3.2 Smart Maintenance Mode of Utility Tunnel Based on BIM

Using the existing BIM model built in the design and construction phase as a foundation, the maintenance information acquired by installing sensors and monitoring equipment can be integrated with the existing BIM model, which prepares the fundamental platform for smart maintenance of the utility tunnels [11, 12]. The core contents of the utility tunnel maintenance system are composed of building parameters, location parameters, product parameters, functional parameters, and management parameters. The main functions of the system include information query, intelligent control of pipe line net, safety pre-warning, maintenance for the record and rescue plan simulation. The composition of the smart whole maintenance system is as shown in Fig. 3.

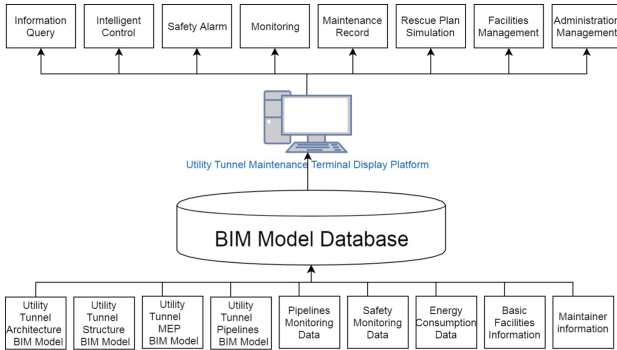


Fig. 3. Design thought of integrated pipe gallery operation and maintenance management system based on BIM

4 Conclusion

In order to improve the efficiency of design and construction integration of the conventional utility tunnel, by studying the problems existing in the traditional design process and construction management mode of the utility tunnel, the design and construction integrated optimization model based on BIM is established. On the other hand, in order to solve the problems such as low efficiency and management difficulties existing in the traditional maintenance mode of the utility tunnel, a smart operation and management mode based on BIM technology was established through the study on the application of BIM technology to the operation and maintenance management system of the utility tunnel. Combining the integrated optimization mode of design and construction with the smart maintenance mode, the visualization, integration, scientific and intelligent management in the whole life cycle of urban utility tunnel can realize in some degree.

References

1. Bai, H.L.: A trend study of urban common tunnel development. *Chin. Municipal Eng.* **6**, 78–81 (2015)
2. Ministry of Housing and Urban-Rural Development, P.R.C. Technical code for urban utility tunnel engineering: GB 50838-2015[S]. China Planning Press, Beijing (2015)
3. Chen, J., Wu, D.K., Ren, J.B., Li, J., Liu, Q.B., et al.: Using cloud-BIM to improve the collaborative design of construction projects. *J. Eng. Manage.* **5**, 27–31 (2014)
4. Lee, H., Kim, J., Banerjee, A.: Collaborative intelligent CAD framework incorporating design history tracking algorithm. *Comput. Aided Des.* **42**(12), 1125–1142 (2010)
5. Wang, J.X., Tang, M.X., Song, L.N., et al.: Design and implementation of an agent-based collaborative product design system. *Comput. Ind.* **60**(7), 520–535 (2009)
6. Man, Q.P., Li, X.D.: Collaborative construction based on ubiquitous computing and BIM. *Chin. Civ. Eng. J.* **s2**, 311–315 (2012)

7. Chen, Y.G., Ding, J.X.: Research on collaboration mechanism of design-construction integration for integrated pipe gallery based on BIM technology. *J. Inf. Technol. Civ. Eng. Archit.* **4**, 56–63 (2018)
8. Hai, L.: Design, scheme brief of utility tunnel operation and maintenance system based on BIM technology. *Municipal Technol.* **5**, 167–168 (2018)
9. Li, Q., Xu, G.Q., Wei, H.M.: Research on the operations management system of utility tunnel based on BIM. *Chin. J. Undergr. Space Eng.* **4**, 287–292 (2018)
10. Li, C.M.: Life cycle management: the Singapore model of underground integrated pipe gallery. *Chin. Explor. Des.* **3**, 72–75 (2016)
11. Zheng, F.S., Tao, W.X., Pan, L.B., et al.: Research on the intelligent management platform of urban underground pipelines. *Chin. J. Undergr. Space Eng.* **11**(Supp. 2), 378–382 (2015)
12. Yang, D.F., Liu, X.D., Su, F., et al.: Research and Application of intelligent operation management of urban utility tunnel. *J. Inf. Technol. Civ. Eng. Archit.* **9**(6), 28–33 (2017)



A Review of Air Quality Monitoring System Based on Crowdsensing

Min Huang^(✉) and Xia Wu^(✉)

Department of Software Engineering,
The South China University of Technology, Guangzhou, China
minh@scut.edu.cn, wuxia_email@163.com

Abstract. In the process of modern intelligent cities, how to effectively collect data on air pollution is a very important issue for citizens. Nowadays, more sensors deployed on mobile terminals, and lots of users use these mobile devices to cooperate with the Internet to realize sensing task distribution and sensing data collection, so as to complete large-scale and complex social awareness tasks, generally, this process is called crowdsensing. The traditional monitoring method is to use fixed base stations, which lacks flexibility and precision. On the other hand, crowdsensing provides new ideas for solving those problems, and more people can participate in it. So far, many studies have proposed different systems to monitor air quality. We review these articles and categorize the systems, then, we compare these systems and analyze their disadvantages and disadvantages, similarities and differences. Finally, we present the challenges of research to provide a reference for new researchers entering the field.

Keywords: Air quality monitoring system · Crowdsensing · Sensor

1 Introduction

Crowdsensing combined crowdsourcing ideas and mobile device sensing to obtain data and is a form of expression of the Internet of Things (IoT) and service-oriented computing [1]. Crowdsensing helps ordinary people or professionals to collect data, analyze information, and share knowledge by forming interactive, participatory networks of sensing through existing mobile devices and publishing sensing tasks to individuals or groups in the network [2]. Participatory sensing [3], mobile crowdsensing [4], people-centric sensing [5], social sensing [6], even though, these concept focuses differently, but it is a way for people to assist mobile devices in data collection, information sorting and knowledge discovery. At present, there is no clear definition of crowdsensing. As shown in Fig. 1, the crowdsensing process basically includes five different tasks [7]: (1) sampling process; (2) filtering process; (3) data transfer; (4) data processing; and (5) results presentation. All of these tasks can be done by different hardware components.

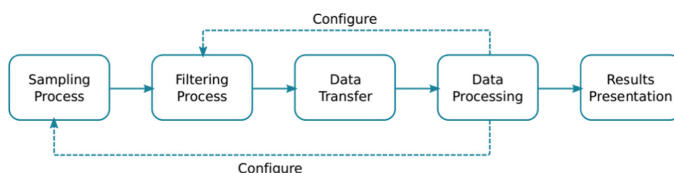


Fig. 1. [7] Crowdsensing steps

Since the industrial revolution, numerous factories and vehicles have been emitting large amounts of exhaust and air pollutants into the atmosphere. These air pollutants are harmful to humans, animals and ecosystems. People are paying more attention to environmental protection and health. Shao et al. [8] constructed a screening method for atmospheric health benchmark target pollutants through pollutant detection rate score, toxic effect score and population exposure score. This method is used to screen a list of candidates for atmospheric health benchmark target pollutants. Some studies to explore the relationship between major pollutants in the atmosphere, which will help the relevant departments to take targeted measures. Wang et al. [9] investigated the status of air pollution in 16 cities, and used some special software to analyze and calculate the data of six major atmospheric pollutants such as SO_2 , NO_2 , PM_{10} , $\text{PM}_{2.5}$, CO and O_3 by principal component analysis. To measure the air quality, we usually use the air quality indicator (AQI) defined by the US Environmental Protection Agency (EPA) [10].

At present, crowdsensing provides a new model for sensing the environment. Crowdsensing has many advantages such as flexible deployment economy, multi-source heterogeneity of sensing data, wide coverage and high expansion and versatility.

Because of these advantages, research on air quality monitoring has emerged in an endless stream. For example, Wang et al. [11] deploy WSN to monitor the concentration of CO and PM contaminants. Penza et al. [12] install gas sensors at certain locations in the city to measure changes in CO , H_2S (hydrogen sulfide), NO_2 and SO_2 . The collected data is converted into a data quality target format as defined by the European Directive [13]. Brienza et al. [14] develop a sensor kit that allows people to easily install gas sensors at home, monitor community air quality, and share monitoring data via social networks. Most mobile phones currently do not have embedded gas sensors, so the research on air quality perceived by mobile crowdsensing generally selects external sensing devices. And more related application cases will be discuss in the Sect. 2.

The paper summarizes the mainstream research on air quality monitoring system based on crowdsensing. These studies are divided into commonly used systems according to the way of data collection, such as: mobile phone camera, wearable, bicycle, etc. First, we review these systems, and then we compare these systems, analyze their disadvantages and disadvantages, similarities and differences. At last, we propose the challenges and possible research directions in the field of to provide reference for scholars. The paper is organized as follows. Different data collection system is discussed in Sect. 2. Research challenges are discussed in Sect. 3. Finally, Sect. 4 concludes.

2 Different Crowdsensing Systems for Air Quality Monitoring

2.1 Camera-Based Systems

As technology continues to evolve, photographs taken with cameras can be used to determine current partial environmental pollutant indices.

Liu et al. [15] use the camera in the phone to take a picture to estimate the value of $PM_{2.5}$. They proposed a learning-based (LB) method to accurately infer the $PM_{2.5}$ concentration using readings from nearby reference sensors to train the predictive model. Figure 2 shows the architecture of their proposed system. Fixed $PM_{2.5}$ sensors will also be connected to the server to improve data accuracy. AirTick [16] is a mobile app that turns any camera-enabled smart mobile device into an air quality sensor for a variety of air quality. AirTick uses image analysis and deep learning techniques to generate accurate estimates of air quality. Liu et al. [17] introduce the mobile Third-Eye application, which can also turn a mobile phone into a high-quality $PM_{2.5}$ monitor. They explore cluster search and web crawling to efficiently build large data sets, which include weather data and air pollution data for outdoor images taken by mobile phones.

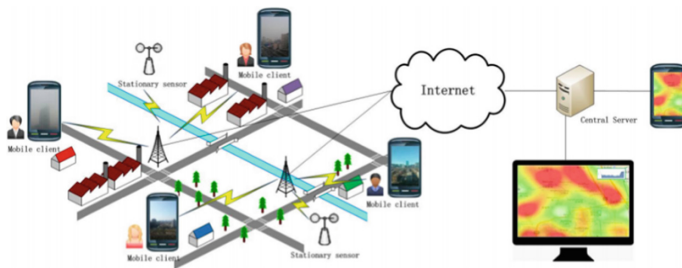


Fig. 2. [15] System architecture

Table 1. Comparison of application cases of camera-based systems above

References	[15] (2015)	AirTick [16] (2017)	Third-Eye [17] (2018)
Sensors	Camera	Camera	Camera Monitoring station
Main technology	Image Processing LB approach	Image analysis Deeplearning	Web crawler, Deeplearning (CNN, LSTM)
Monitoring object	$PM_{2.5}$	$PM_{2.5}$	$PM_{2.5}$, CO, NO_2 , O_3 , SO_2
Advantage	High precision	Extract source information for renewable energy	Combined with other gases, accuracy is higher than AirTick, portability and price
Disadvantage	The system is not open, incentives, user privacy	Insufficient precision, incentive mechanism	Insufficient precision at night, excessive time density, not applicable indoors
Usefulness evaluation	90% accuracy	96% of respondents think it is useful	Classification accuracy 81.55%

Comparing the above methods, we can see that the structure of these systems is similar to that of Fig. 2, the difference lies in the technical processing of these pictures, which include combining monitoring data on the network to verify the sensor, fine-grained monitoring range to improve accuracy, etc. Finally, the value of the $PM_{2.5}$ is displayed in numerical form in conjunction with the map. For those specific systems design, you can refer to the original articles.

Table 1 shows the differences and advantages of these systems. We can see that advantage is that it does not require additional external sensors, is cheap, and easy to obtain data. You can check the concentration of the current location at any time. Moreover, it is possible to know the concentration in other places with the participation of most people, so as to choose a better travel plan. The main disadvantage is that the accuracy at night is not high, and the main monitoring level is the concentration of $PM_{2.5}$, and other invisible gases cannot be directly monitored.

2.2 Wearable-Based Systems

Wearable devices, such as wrist straps, attached to belts or backpacks, will integrate commercial and off-the-shelf gas sensors for personal air pollution sensing. However, previous studies lacked a comprehensive survey of the accuracy of air pollution sensing for wearable devices.

Table 2. Comparison of application cases of wearable-based systems above.

References	[18] (2013)	[19] (2015)	W-Air [20] (2018)
Sensors	CO, NO ₂ , SO ₂ , temperature, relative humidity, atmospheric pressure	CO, SO ₂ , NO ₂ , temperature, relative humidity, atmospheric pressure	CO ₂ (indoor) O ₃ (outdoor), VOC, temperature
Main technology	OpenIoT, middleware	Complete electrical, mechanical, and software design	Internet of Things, neural network, semi-supervised learning
Communication	Bluetooth	Bluetooth	Bluetooth
Advantage	Periodic sensing, high data collection flexibility, real-time alarm prompt	Prove the feasibility of the program	Mainly biased towards sensor calibration, exclude human interference, accurate gas concentration detection
Disadvantage	Less participants	Experimental condition is in the laboratory	Environment, activity, energy, delay
Usage time	3 days	65 h	5 h

Antonic et al. [18] introduce an urban crowdsensing aware application that monitors air quality through the use of specially designed wearable sensors and mobile phones. The application is based on the OpenIoT platform and uses sensors to generate information about pollutant gas concentrations and meteorological conditions. Oletic et al. [19] introduce the design of a battery-powered wearable sensor node that includes two electrochemical gas sensors, temperature, relative humidity and atmospheric pressure sensors, and a Bluetooth connection. W-Air [20] is an accurate wearable device personal multi-contamination monitoring platform that uses a sensor fusion calibration scheme to recover high quality environmental pollutant concentrations from human interference. It also utilizes a neural network with shared hidden layers to improve calibration parameter training with fewer measurements and utilizes semi-supervised regression for calibration parameter updates for less user intervention. Table 2 shows the comparison of these types of wearable devices. For those specific systems design, you can refer to the original articles.

Generally, the architecture of the wearable-based systems are similar, and the mobile device communicates with the external sensor through Bluetooth. The difficulty of this system is how to eliminate the interference of the user's wearing position and different user actions. The advantage of this detection solution over camera-based system is that more sensors can be used to monitor different gases in order to take into account the air quality. Many current wearable devices can measure human heart rate, user pace, sleep quality and other human health indicators. In the future, if the gas sensor becomes smaller and integrated into these wearable devices, users can easily know the surrounding air quality and their health.

2.3 Bicycle-Based Systems

Mobility and environmental conditions are key factors in the urban environment, affecting well-being and quality of life. In this case, sensors, smart mobility, networks and connections can play an important strategic role and are used to improve the data and information available to the public administration and to every citizen. In this way, they can get more support for sustainable and conscious behavior and get useful information and services to improve their daily activities.

BikeNet [21] is a scalable mobile sensing system for rider experience maps that leverages the principles and technologies of opportunistic sensor networks. BikeNet represents a multi-faceted sensing system that uses a network of bicycles. The SmartBike [22] platform provides: real-time remote location of the user's bicycle, anti-theft service, information about the route and air pollution monitoring. Figure 3 shows the logic architecture of the SmartBike. Canarin II [23], a smart electric bicycle prototype, was equipped with a system for collecting data on particulate matter and elaborated and shared the data. Contaminant sensors can transmit these data over heterogeneous networks via the city's wireless network or through passenger smartphones.

Table 3 shows the comparison of these types of bicycle devices. The advantage of the bicycle-based system over wear is that these sensors are integrated on the bicycle, do not require the user to carry it, and the bicycle is a green way of travel without the discharge of contaminants. For example, many cities provide urban bicycles. If these

bicycles integrate sensors that monitor air quality, combined with GPS, heart rate and other sensors, the collected data will be transmitted to the mobile phone or directly to the server. Users can view the current status on the client anytime, anywhere. They can also know their own health, etc., and users participate in this activity to check the surrounding pollution before traveling, and you can also choose the path of travel.

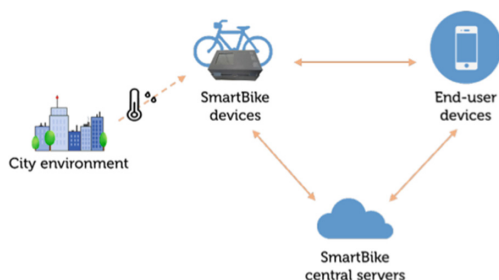


Fig. 3. [22] High level architecture design of the SmartBike platform

Table 3. Comparison of application cases of bicycle-based systems above

References	[21] (2009)	SmartBike [22] (2017)	Canarin II [23] (2018)
Collected data	Heart rate, galvanic skin response, Self-built bicycle area network	CO, temperature, humidity and pressure, position information, monitoring information	Time, GPS coordinates, temperature, humidity, pressure, formaldehyde, PM ₁ , PM _{2.5} , and PM ₁₀
Network	Image Processing LB approach	Bicycle network	Vehicle network
Communication	GSM,GPRS, Bluetooth	Bluetooth	Wi-Fi, Mobile communication
Advantage	Support real-time sensing and delay tolerance, localized services, and good system portability	Survey to collect user needs, high accuracy, real-time location information, anti-theft services	Bicycle travel path, biker’s exercise information, bicycle safety

2.4 Car-Based System

The bicycle-based system used in the front is a slow moving system. A similar mobile-based system is to place the sensor on the car, which moves faster.

Hu et al. [24] develop a mobile-based system by cars to monitor the concentration of CO₂ gas in urban areas. Specifically, each car is equipped with four components:

CO₂ sensor, GPS receiver, GSM module, and Jennic board. By combining the positioning information from the GPS receiver and the sensing data from the CO₂ sensor, the GSM module then periodically transmits the monitoring data to a nearby GSM base station through GSM short messages. Sivaraman et al. [25] propose a HazeWatch. In HazeWatch, the architecture as shown in the Fig. 8, a driver can choose to mount either a cheap but simpler oxide sensor or an expensive. In addition, commercial monitors that provide more accurate contaminant detection are installed on the roadside to calibrate the readings of the car’s sensors. The mode of the car is similar to that of Fig. 4, and the above systems’ comparison are shown in the Table 4. With the improvement of technology, the communication protocol, network bandwidth, type of sensor, type of server used, and accuracy of the sensor are changed. Compared with bicycle systems, this system needs to eliminate the interference of automobile exhaust and pay more attention to air mobility. Generally, the vehicle wireless sensor network is combined with the mobile phone GSM, 3G and other communication methods to transmit the data of the polluted gas concentration to the mobile device, then transmit to the server through the mobile device, and the result is presented to the user in a map.

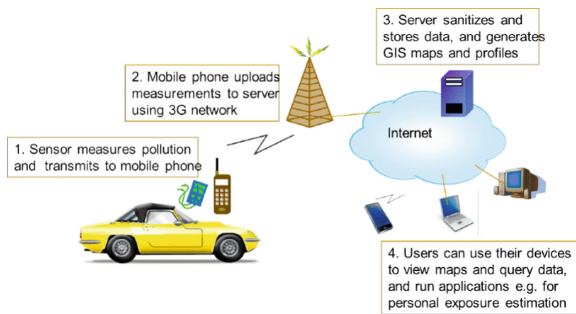


Fig. 4. [25] HazeWatch’s system architecture

Table 4. Comparison of application cases of car-based systems above

References	[24] (2010)	HazeWatch [25] (2013)
Collected data	CO ₂ , GPS	O ₃ , NO ₂ , CO
Network	Vehicle wireless sensor network	Bicycle network
Communication	GSM, zigbee	3G, mobile phone
Advantage	Mobility	Personal pollution map, path planning, real-time tag upload data

2.5 Discussion

In addition, some papers mainly study air quality information shared by users in social networks to estimate air quality. But the method has a large degree of error, and the authenticity of information is difficult to judge. Another, in large industrial areas or rural areas where transportation facilities are poor or non-existent, some studies propose Unmanned Aerial Vehicles with pollution sensors to monitoring air quality.

Handheld-based system were introduced more than a decade ago, and Common Sense [26] allow users to view current ozone concentrations in real time. Citsense [27] can monitor NO₂, O₃, CO, temperature, humidity and pressure indicators. AirSense [28] is an opportunistic crowdsensing based air quality monitoring system designed to collect and aggregate sensor data to monitor air pollution in nearby and urban areas. You can read the original papers to know how they work. Moreover, many other architectures are similar to AirSense, and other applications change the method implemented in each step, so as to realize the idea of crowdsensing more efficiently and conveniently.

However, the common system mentioned above is to use sensors to collect data and build an overall system architecture according to different acquisition methods. These architectures are similar, which are taken from different articles above. The users' mobile devices and the sensors collect data together, communicate via Bluetooth or other communication protocols, and then transmit the data to mobile devices. The data may be processed in the mobile phone, or transmitted to the server background through Wi-Fi, 3G, 4G, etc. to process, and finally generate a user-friendly interface to display the result. Users can view the air quality of the user's location and the air quality of other users who have participated in or not participated in the event through internet access devices. Table 5 shows the commonalities and differences between these systems.

Table 5. Comparison of all systems above

System	Common sensors	Advantage	Disadvantage
Camera	Camera	No need for other sensors, convenient	Only monitor visible substances, accuracy
Wearable	PM _x , CO, CO ₂ , NO ₂ , O ₃ , temperature, humidity, atmospheric pressure, etc.	High portability	Subject to human action, skin, textile interference
Bicycle		Green travel, city bike or custom bike available	Air flow interference
Car		Travel path recommendation	Air flow, vehicle exhaust interference

3 Challenge

There are many applications of crowdsensing in air quality monitoring, most of which are realized by combining mobile sensor technology, data mining technology, machine learning, image processing technology, etc. The challenges encountered include the challenges of these technologies themselves. Overall challenges include, for example, we need to consider user participation issues, hardware devices and software design issues. The paper will present the challenges of technology application from the five steps of crowdsensing processing.

- (1) Sampling process. The main challenges are the error of the sensor itself, sensor power consumption, volume, and sensor calibration methods. Sensors from different manufacturers have different errors but the general error is an acceptable range. The power consumption, size and cost of the sensor will affect the range of use.
- (2) Filtering process. The collected data may repeat too much useless data. For example, the data reference in the same location for a period of time is similar and does not require full transmission. In addition, the collected data may be incorrect and needs to be filtered.
- (3) Data transfer. Data transmission mainly solves the balance between real-time and opportunistic. The problem of large data volume also needs to be considered. Data compression technology can be adopted to consider the bandwidth and delay. There are some problems with the format of data uploading and the data transmission methods.
- (4) Data processing. At this stage, most of the research is to put the background data in the cloud, and the data processing generally refers to the interpolation technique used to reconstruct the pollution map.
- (5) Results presentation. The result display is the way the results are presented to the system administrator. The most useful representation is the graphical mapping of the target area. The speed at which the structure is presented is a challenge.

4 Conclusion

Air pollution is a global concern that affects everyone's health. As a result, citizens are increasingly concerned about environmental monitoring, but traditional air quality monitoring is through fixed base stations, which are not only costly but also less flexible. The emergence of crowdsensing can solve the above problems well. Combined with the power of the public, it can obtain air pollution data in real time and can contribute to air quality monitoring. In this paper, we investigate various common air quality monitoring technologies, from data collection systems such as camera-based, wearable-based, bicycles, cars. In addition, we compare those systems and sum up the difficulties of its research to provide reference for scholars and to help them quickly know the research field and find research problems, and contribute to air quality monitoring public utilities.

Acknowledgments. The project is supported by Guangdong province science and technology planning projects (No. 2016B070704010), and Guangdong province science and technology planning projects (NO. 2016B010124010).

References

1. Huang, M., Chen, Y.: Internet of things and intelligent devices and services. *CAAI Trans. Intell. Technol.* **3**(2), 73–74 (2018)
2. Huang, M., Bai, Y., Chen, Y., Sun, B.: A distributed proactive service framework for crowd-sensing process. In: *Proceedings of IEEE Conference 13th International Symposium on Autonomous Decentralized Systems*, pp. 68–74 (2017)
3. Estrin, D.: Participatory sensing: applications and architecture. *IEEE Internet Comput.* **1**, 12–42 (2010)
4. Ganti, R.K., Ye, F., Lei, H.: Mobile crowdsensing: current state and future challenges. *IEEE Commun. Mag.* **11**, 32–39 (2011)
5. Zhao, D., Ma, H., Liu, L.: Energy-efficient opportunistic coverage for people-centric urban sensing. *Wirel. Netw.* (2014)
6. Madan, A., Cebrian, M., Lazer, D., Pentland, A.: Social sensing for epidemiological behavior change. In: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, pp. 291 – 300. ACM Press (2010)
7. Alvear, O., Calafate, C., Cano, J.C., et al.: Crowdsensing in smart cities: overview, platforms, and environment sensing issues. *Sensors* **2**, 460 (2018)
8. Shao, T., Miao, X., Zhou, Z., et al.: A screening method for the target pollutants of atmospheric health benchmark in China. *Environ. Chem.* (2017)
9. Wang, X.L., Liu, C., Guan, W.L.: Study on the relationship of city primary air pollutants based on principal component analysis method. *J. Tianjin Univ. Technol.* (2015)
10. Schwanke, R.C., Marcon, R., Bento, A.F., et al.: EPA- and DHA-derived resolvins' actions in inflammatory bowel disease. *Eur. J. Pharmacol.* **785**, 156–164 (2016)
11. Wang, C.H., Huang, Y.K., Zheng, X.Y., Lin, T.S., Chuang, C.L., Jiang, J.A.: A self-sustainable air quality monitoring system using WSN. In: *Proceedings of IEEE International Conference Service-Oriented Computing and Applications*, pp. 1–6 (2012)
12. Penza, M., Suriano, D., Villani, D., Spinelle, L., Gerboles, M.: Towards air quality indices in smart cities by calibrated low-cost sensors applied to networks. In: *Proceedings IEEE SENSORS*, pp. 2012–2017 (2014)
13. European Commission, “Air quality–existing legislation” (2016). <http://ec.europa.eu/environment/air/quality/legislation/existingleg.htm>
14. Brienza, S., Galli, A., Anastasi, G., Bruschi, P.: A low-cost sensing system for cooperative air quality monitoring in urban areas. *Sensors* **15**(6), 12242–12259 (2015)
15. Liu, X., et al.: PM2.5 monitoring using images from smartphones in participatory sensing. In: *Computer Communications Workshops IEEE*, pp. 630–635 (2015)
16. Pan, Z., Yu, H., Miao, C.: Crowdsensing Air Quality with Camera-Enabled Mobile Devices (2017)
17. Liu, L., Liu, W., Zheng, Y., Ma, H., Zhang, C.: Third-Eye: A Mobilephone-Enabled Crowdsensing System for Air Quality Monitoring (2018)
18. Antonic, A., Bilas, V., Marjanovic, M., et al.: Urban crowd sensing demonstrator: sense the Zagreb air. In: *International Conference on Software, Telecommunications and Computer Networks*, pp. 423–424. IEEE (2014)

19. Oletic, D., Bilas, V.: "Design of sensor node for air quality crowdsensing. In: Sensors Applications Symposium, pp. 1–5. IEEE (2015)
20. Maag, B., Zhou, Z., Thiele, L.: W-Air: Enabling Personal Air Pollution Monitoring on Wearables (2018)
21. Eisenman, S.B., Miluzzo, E., Lane, N.D., et al.: BikeNet: a mobile sensing system for cyclist experience mapping. *ACM Trans. Sens. Netw.* **6**, 1–39 (2010)
22. Corno, F., Montanaro, T., Migliore, C., et al.: SmartBike: an IoT crowd sensing platform for monitoring city air pollution. *Int. J. Electr. Comput. Eng.* **6** (2017)
23. Aguiari, D., Delnevo, G., Monti, L., et al.: Canarin II: designing a smart e-bike eco-system. In: IEEE Consumer Communications & NETWORKING Conference 2018, pp. 1–6. IEEE (2018)
24. Hu, S.C., Wang, Y.C., Huang, C.Y., Tseng, Y.C., Kuo, L.C., Chen, C.Y.: Vehicular sensing system for CO₂ monitoring applications. In: Proceedings IEEE Asia Pacific Wireless Communication Symposium, pp. 168–171 (2009)
25. Sivaraman, V., Carrapetta, J., Hu, K., Luxan, B.G.: HazeWatch: a participatory sensor system for monitoring air pollution in Sydney. In: Proceedings of IEEE Conference Local Computer Networks, pp. 56–64 (2013)
26. Dutta, P., et al.: Common sense: participatory urban sensing using a network of handheld air quality monitors. In: International Conference on Embedded Networked Sensor Systems, SENSYS 2009, Berkeley, California, USA. DBLP, pp. 349–350, November 2009
27. Ziftci, C., et al.: Citisense: mobile air quality sensing for individuals and communities design and deployment of the Citisense mobile air-quality system. In: International Conference on Pervasive Computing Technologies for Healthcare IEEE, pp. 23–24 (2012)
28. Dutta, J., Gazi, F., Roy, S., et al.: AirSense: opportunistic crowd-sensing based air quality monitoring system for smart city. In: Sensors, pp. 1–3. IEEE (2017)



Correction to: An Improved Vehicle Detection and Tracking Model

Libin Hu, Zhongtao Li, Hao Xu, and Bei Fang

Correction to:

Chapter “An Improved Vehicle Detection and Tracking Model” in: X. Zeng et al. (Eds.): *International Symposium for Intelligent Transportation and Smart City (ITASC) 2019 Proceedings*, SIST 127,
https://doi.org/10.1007/978-981-13-7542-2_8

In the original version of the book, belated affiliation corrections in the chapter “An Improved Vehicle Detection and Tracking Model” for all the authors have been incorporated as indicated below:

1. School of Information Science and Engineering, University of Jinan, Jinan 250022, China
2. Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan 250022, China

The updated version of this chapter can be found at
https://doi.org/10.1007/978-981-13-7542-2_8

© Springer Nature Singapore Pte Ltd. 2019
X. Zeng et al. (Eds.): ITASC 2019, SIST 127, p. C1, 2019.
https://doi.org/10.1007/978-981-13-7542-2_29

Author Index

B

Bai, Yan-feng, 206
Bao, Li-xia, 241
Bi, Yunrui, 107

C

Cao, Ketu, 63
Chen, Yining, 138, 197
Chen, Yinong, 13, 232
Chen, Yujia, 145
Cui, Ke, 155

F

Fan, Ding, 94
Fang, Bei, 84
Feng, Tao, 49

G

Gai, Songxue, 127, 186
Guo, Ying, 232

H

Hao, Wu, 94
He, Chang-xuan, 206
He, Junxiang, 177
Hu, Chuanpeng, 277
Hu, Libin, 84
Huang, Jicheng, 138
Huang, Min, 286

L

Li, Jinping, 251
Li, X. H., 74

Li, Xinhan, 251
Li, Zhongtao, 84
Liu, Liqun, 265
Lu, Mengdan, 107
Luo, Jing, 13

M

Ma, Yan, 107
Ma, Zhongzheng, 138, 197

N

Nian, Guangyue, 251
Nie, Dongqing, 222

O

Ou, Dongxiu, 155

P

Pan, Xiaofeng, 49
Peng, Zhang, 94

S

Sun, Daniel(Jian), 13, 251
Sun, X. H., 74
Sun, Xiaohui, 63

T

Tang, Maojie, 155

W

Wang, Gang, 107, 214
Wang, L. X., 74
Wang, Lixiao, 49, 63

Wang, Qiulan, 39
Wang, Sayi, 39
Wang, Yizeng, 1, 138, 177, 197, 265
Wang, Yuan-qing, 241
Wu, Chaoyang, 28, 127, 167
Wu, Suhua, 107
Wu, Xia, 286

X

Xiaobo, Wen, 94
Xiong, Qipeng, 1, 28, 167, 265
Xu, Hao, 84
Xu, Kui, 107
Xu, Xinchun, 1, 177, 265

Y

Ye, Nixuan, 138, 197
Ying, Peiran, 177
Yuan, Tengfei, 145, 167

Z

Zeng, Xiaoqing, 1, 13, 28, 127, 138, 145, 167,
177, 186, 197, 265
Zhan, Jifei, 28, 127
Zhang, Chi, 118
Zhang, Shilang, 277
Zheng, Lianzhen, 232
Zhu, Guohua, 118
Zuo, Zhi, 49, 63, 74